AD_____

Award Number: W81XWH-05-1-0397

TITLE: BioTerNet Networking and Strain Tracking

PRINCIPAL INVESTIGATOR: Myron Sasser, Ph.D.

CONTRACTING ORGANIZATION: MIDI, Incorporated
Newark, DE 19713-1148

REPORT DATE: July 2006

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-072006	Final	1 Aug 2005 – 31 Jul 2006

4. TITLE AND SUBTITLE

BioTerNet Networking and Strain Tracking

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0397

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Myron Sasser, Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

MIDI, Incorporated
Newark, DE 19713-1148

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

TheobjectiveofBioTerNetwastodevelopasystem,basedonMIDI'sSherlockfattyacidtechnology,toaidintheidentificationofabacterialbioterrorismevent andassistinconsequencemanagement.Thisprojectwasconceivedasa36montheffortwithfourkeytechnicalobjectives:[1](months1–14)Developafastbacterialfattyacidextractionprocedurethatminimizesvariabilityduetooperatorandlabconditionstoimprovestraintrackingcapability.Provideresults24hourssoonerbyanalyzingsinglecolonies.Improvesignalprocessingalgorithmsforbetterstraintracking.[2](months14–30)Createacomprehensivelibraryofbacteriathatworkswiththefastextractionprocedure.[3](months1–25)EnhanceSherlockfromastandaloneapplicationtoasecureclient-serverapplication.[4](months21–28)Developalgorithmstomonitorcentralizeddatabaseforsuspiciousevents.(months30–36)Testdeployment.Havinglearnedearlythatonlythefirstyearwouldbefunded,theprojectfocusedonobjectivesthatwouldmaximizethedeliveredvalue bydevelopingthefastextractionprocedureandcollectingasmuchdataaspossibletocreatethelibrarynamedin[2]above.Afastextractionprocedurethatimprovessensitivityandminimizesvariabilitywasdeveloped.Abacteriallibrarywascreatedusingsoftwaretechniquestoreusedataacquiredusingan oldersampleprepprocedure.Enhancedtracking/clusteringwasaddedtoSherlock.Aprototypedemonstratingthefeasibilityofnetworkingandacentral datastorewascreated.

15. SUBJECT TERMS
bacteria,biodefense,bioterrorism,identification,networking,strainidentification,clustering,networking,fattyacids

16. SECURITY CLASSIFICATION OF:				18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	39	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Introduction

Bacterial agents used in biowarfare/bioterrorism may consist of well-known, characterized agents, or unknown and possibly genetically engineered bacteria. In order to save the maximum number of lives and to allow for effective countermeasures and investigative activities after a bacterial agent release, it is crucial to quickly identify the specific bacterial strain and source used in an attack.

A basic capability to address this need is already present in the existing MIDI Sherlock System. The Sherlock System uses a gas chromatograph to analyze extracted bacterial organisms for their inherent cellular fatty acid (CFA) components. The fatty acid profile is compared to a library of known bacterial entries to determine the species identification for the sample. Sherlock has library entries for all of the major bacterial bioterrorism agents as well as routine bacterial organisms. Sherlock can compare bacterial samples to each other. It is a single system that can both identify known organisms and help recognize repeat occurrences of known or unknown strains. Sherlock is a strong platform for building a system to use for epidemiological surveillance and recognition of bioterrorism/biowarfare attacks. To meet these requirements, the following need to be developed:

- A fast extraction procedure that minimizes variability in the fatty acid profile due to operator and laboratory conditions for faster results and higher confidence in recognizing repeat occurrences
- Increased sensitivity of the extraction procedure and gas chromatography method with the goal of requiring only a single colony of bacteria to eliminate 24 hour growth requirement
- A bacterial library that works with the fast extraction
- The ability to process data across multiple labs to speed recognition of events spread over wide areas

Body

The scope of work for the project spreads over a 36-month period as setout in the proposal's statement of work. Early in the first year it was learned that funding for the second and third years would not be available. In order to deliver a more functional result at the end of the first year, the priority and approach for creating a comprehensive library of bacteria (Technical Objective 2 below) were modified to build a workable library during the first year.

Statement of Work (from Proposal)

The overall objectives of Sherlock BioTerNet are to aid in the identification of a covert bacterial bioterrorism or biowarfare event and to assist in consequence management of an overt event. Sherlock BioTerNet is a system that identifies species and tracks strains of bacteria within an automated networking environment. The expected results of BioTerNet include: developing a novel sample preparation protocol with increased sensitivity, improved reproducibility and reduced labor costs; implementing BioTerNet, a client-server application containing a relational database that will provide automated monitoring for strain matching of select agents and other organisms of interest; and incorporating in BioTerNet a rapid response capability in the case of discovery of new infectious agents, achieved through electronic addition to the libraries used by BioTerNet.

Technical Objective 1: Develop a rapid procedure for extraction of bacterial fatty acids that minimizes variability due to the operator and to laboratory conditions, enhancing the strain tracking capability of the system. Provide results up to 24 hours faster than the current system by analyzing single colonies from the primary isolation plate.

- months 1-8: Develop sensitivity modifications for single colony analysis.
- months 9-14: Develop signal processing and data normalization algorithms to improve reliable detection of repeat occurrences of an organism.
- months 3-11: Develop a rapid extraction technique and evaluate for single colony analysis.

Technical Objective 2: Create a comprehensive library of bacteria for identification using the protocols developed in Technical Objective 1, so that identification can be performed using the new extraction protocol.

- months 14-25: Library creation and internal validations to assess the discriminatory ability of the library entries.
- months 26-28: Update signal processing algorithms.
- months 28-30: Validation studies coordinated with regulatory agencies.

Technical Objective 3: Enhance Sherlock from a standalone application with a proprietary data store to a secure client-server application with an underlying relational database, allowing Sherlock BioTerNet to rapidly track bacterial strains across multiple sites in real time.

- months 1-14: Replace Sherlock's proprietary file system with a relational database.
- months 13-25: Create client-server version of Sherlock.
- months 14-20: Add automated tracking capability to Sherlock BioTerNet.

Technical Objective 4: Develop algorithms that automatically monitor BioTerNet's centralized database for suspicious events that may be indicative of a bioterrorism/biowarfare event.

- months 21-23: Design Suspicious Event Detection Module.
- months 24-28: Implement and test Suspicious Event Detection Module.

Months 30-36: Sherlock BioTerNet will be deployed and tested at multiple sites to verify its performance.

Having learned early that only the first year would be funded, the project focused on objectives that would maximize the delivered value in the first year. Specifically the project concentrated on improving the speed and accuracy of recognizing a bacterial based bioterrorism/biowarfare event and identifying the bacterial agent involved. Accomplishing these goals requires:

- Development of the fast, less variable extraction procedure described in Objective 1
- Creation of a workable library for identifying bacteria prepared with the fast extraction procedure sooner than planned in Objective 2. The approach was to develop a software/algorithmic transform that maps data taken with the current (slow) extraction procedure to approximate data from the fast extraction.
- Enhance the library by acquiring cultures of species currently missing from the library, analyzing them and adding their fatty acid profiles to the library.
- Improve the clustering and strain tracking capabilities in the standalone version of Sherlock.
- Scale back implementation of the relational database and client-server version of Sherlock listed in Objective 3 and concentrate on a database design and working client-server prototype to demonstrate feasibility of a networked system.

Objective 1

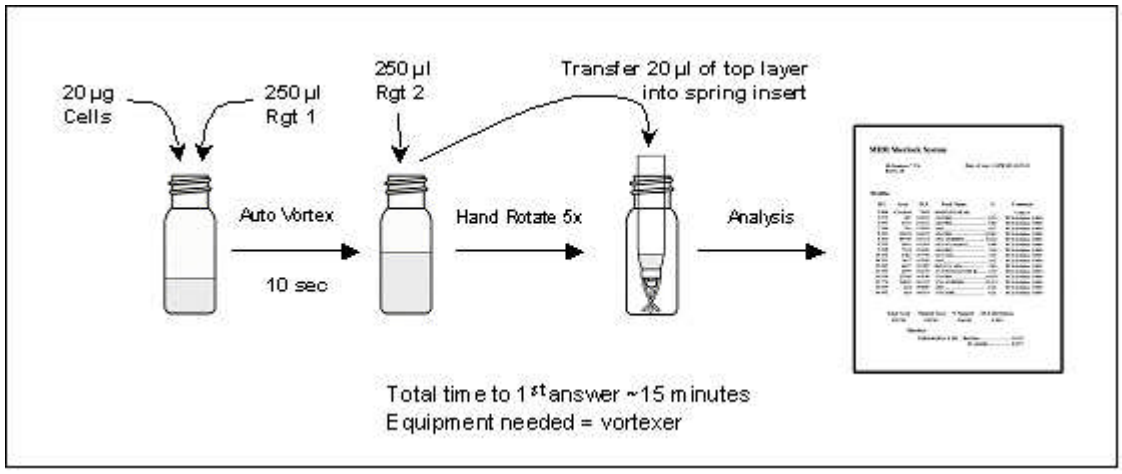
Develop a rapid procedure for extraction of bacterial fatty acids that minimizes variability due to the operator and to laboratory conditions, enhancing the strain tracking capability of the system. Provide results up to 24 hours faster than the current system by analyzing single colonies from the primary isolation plate.

The goal of the "FAST" extraction method is to develop an extraction/derivatization procedure to replace the "Standard" Sherlock method of creation of fatty acid methyl esters (FAMEs) from bacterial cells. The Standard method requires two water baths, one at 100C and one at 80C, uses a vortexer, a clinical rotator, four rather large reagent dispenser bottles, glass Pasteur pipettes, and an ambient temperature water bath. The total volume of reagents used is 8ml, with about 7 ml becoming waste. The desired configuration of the FAST extraction procedure would be to have the procedure be very rapid, highly reproducible, require less equipment and be very robust. It would be highly desirable to have the extract match the current database as much as possible to avoid having to re-create the entire dataset, which would make it possible to have a workable library earlier than planned in the original Statement of Work.

The "Standard" method of extraction and derivatization of bacterial fatty acids in a single sample requires a saponification step (30 minutes in a boiling water bath), cooling in ambient water (1 minute) and uncapping of the vials with addition of the second reagent and recapping (1 minute), heating at 80 ± 2C for 10 ± 1 minutes in a second water bath device, cooling in ambient water (1 minutes), uncapping the vial, addition of the third reagent and recapping; tumbling in a clinical rotator for 10 minutes; uncapping the vial, removing the bottom aqueous layer, addition of the fourth reagent and recapping (2 minutes) and rotating for 5 minutes; the top layer is then removed into a gas chromatography (GC) autosampler vial for analysis. Total elapsed time is a minimum of 60 minutes.

A FAST extraction procedure was developed that requires only one minute from time of harvest of the cells to having the extract in an autosampler insert and ready for analysis. After harvesting the bacteria and inserting them into a GC autosampler vial, the steps in the procedure are: add 250 µl of Reagent 1, vortex for 10 seconds (automatic timer on the vortexer); add 250µl of Reagent 2, vortex for 1 second; add 250µl of Reagent 3, gently swirl the vial for 1 second; remove 50µl of the top layer into a GC autosampler vial insert for analysis. Total time elapsed is less than one minute when done by a practiced technician. The equipment required for the FAST extraction/derivatization is a vortexer, 3 small bottle-top dispensers, and a hand-held pipette.

Figure 1: FAST Extraction Procedure



The FAST protocol is extremely robust. Unlike the current extraction procedures, the FAST protocol is independent of reagent temperature. Table 1 shows the temperature independence based on 4 runs of the *Bacillus subtilis* (ATCC 6633) with reagents at room temperature and at 4 degrees C. The average similarity indices of 0.93 and 0.90 at the two temperature points are equivalent for identification purposes. An index of 1.00 would indicate a perfect match to *B. subtilis*. The total responses indicate nearly identical recovery of fatty acids at the two temperatures.

Table 1: Temperature Effect on Total Response and Similarity Index

Temperature	Total Response	Similarity Index
Room (~23 C)	148,570	0.97
Room (~23 C)	188,760	0.93
Room (~23 C)	130,704	0.96
Room (~23 C)	107,384	0.85
Average	143,855	0.93
S.D.	29,756	0.04
4 C	108,211	0.95
4 C	146,652	0.86
4 C	157,486	0.93
4 C	119,825	0.84
Average	133,044	0.90
S.D.	19,837	0.04

The volume of reagent 1 was reduced to 100µl with reagents 2 and 3 remaining at 250µl each and the extraction performed with no difference in the result. A similar reduction of reagent 3 caused no change in the results when compared to the FAST protocol. The reduction of reagent 2 to 100µl resulted in 2.5x increase in sensitivity of the extraction procedure, but had no detrimental effects when compared to the FAST protocol, except that more care is required when removing the top layer for analysis.

The major negative point of the FAST extraction protocol is that it does not reproducibly extract the fatty acids of the Gram-negative bacteria. MIDI software developers performed tests of the current database with the hydroxy acids removed and found that there was little loss of information due to removal of these compounds from the peak-naming table (see results in Tables 3 and 4). The reason the impact is small is likely due to two factors (1) higher variability of these compounds in extracts reduces their information content and (2) there is enough information in the remaining compounds to differentiate the species.

The reduction in scale of reagents in the FAST protocol results in about a 10x increase in sensitivity over the current Standard protocol. To additionally add sensitivity, the method is modified (from the current RAPID) method to double the sensitivity of analysis, which when coupled with the 10x increase gained by FAST results in a 20x increase in sensitivity. Even more sensitivity can be gained by reducing the volume of reagent 2 to 100µl (as explained above – a factor of 2.5x) and by changing the split ratio of the inlet to "splitless" thereby gaining additional 15x increase of sensitivity. Coupled with the 20x of the FAST protocol, this would potentially result in about a 300x increase of sensitivity over the current RAPID method and about 750x over the Standard method. Preliminary tests indicate that this sensitivity level will allow processing of single colonies but additional studies are needed.

Potential problems arising with a 750x increase of sensitivity are contaminants may become a major factor and uniform harvesting of a barely visible quantity of cells may be difficult. MIDI will continue to investigate these two problems without additional funding.

Contamination can be reduced by assuring use of high purity gases for the GC (as are currently recommended for the Standard and RAPID methods) and by use of high purity reagents for the extraction. The reagent problem will be resolved by having MIDI purchase (and where necessary, purify) the complex extraction reagents and make these available for purchase by users. The problem of uniform harvesting is being approached by development of a video that illustrates correct harvesting procedure to maximize uniformity of harvest and software has been developed that will modify the volume of sample injected or change the split ratio of the GC inlet in relation to over-harvesting (under-harvesting is more difficult when the system is already maximized for analytical recovery of compound).

The FAST protocol should provide very rapid and robust FAME extraction/derivatization and the changes in the compound naming table coupled with the re-generated database should result in a very useful tool, likely replacing the current protocols of sample preparation and analysis. Initial reproducibility studies along with the robustness test citied above may eliminate the need to develop additional signal processing algorithms.

Objective 2

Create a comprehensive library of bacteria for identification using the protocols developed in Technical Objective 1, so that identification can be performed using the new extraction protocol.

The FAST extraction procedure has numerous benefits, but not all Fatty Acids are reliably reproduced. In order to use the FAST extraction procedure, a prototype library has been developed using existing data reprocessed without using the unreliable peaks from the standard method. Table 2 lists the peaks that were "zeroed" out of the data.

Table 2: Peaks "Zeroed" in Existing Data to Create the "No OH" Library

8:0 3OH	14:0 2OH
9:0 3OH	15:0 iso 3OH
10:0 2OH	15:0 2OH
10:0 3OH	15:0 3OH
11:0 iso 3OH	16:1 2OH
11:0 2OH	16:0 iso 3OH
11:0 3OH	16:0 2OH
12:0 iso 3OH	16:0 3OH
12:0 2OH	17:0 iso 3OH
12:1 3OH	17:0 2OH
12:0 3OH	17:0 3OH
13:0 iso 3OH	18:1 2OH
13:0 2OH	18:0 2OH
14:0 iso 3OH	18:0 3OH

Zeroing a peak has the following affect. The peak is still named by the system but a percentage of 0.0 is calculated for it regardless of its response. The implication is that these peaks, while still recognized by the system are no longer used in calculations of Euclidean Distance or Similarity Index; it is just as if these peaks did not exist.

It should be remembered that the existing clinical library has been optimized over fifteen years of use. Sample selection and library entry parameterization have been used to improve the results with this library. Thus it is not unexpected that the overall results for the Experimental "No OH" library would be lower than the overall results of the standard library. Tables 3 and 4 summarize the performance of the Experimental "No OH" library. (Note: The number of samples is lower for "No OH" due to a loss in total response caused by removing all OH peaks.)

Table 3: Comparison Summary of Standard Library vs. No OH Library

Library	Standard	No OH
Total Samples	9747	9706
Correct ID	8623	8285
Percent Correct	88.5%	85.4%
False Positives	1	27

Table 4: Performance of Experimental "No OH" Library Summarized by Genus

Genus	Num Samples	Num Correct	Species % Correct	Num False Pos.
achromobacter	43	19	44.2%	1
acidovorax	25	24	96.0%	0
acinetobacter	108	99	91.7%	0
actinobacillus	74	49	66.2%	0
aerococcus	13	13	100.0%	0
aeromonas	131	117	89.3%	0
afipia	17	17	100.0%	0
alcaligenes	67	22	32.8%	1
amycolatopsis	29	29	100.0%	0
aquaspirillum	3	3	100.0%	0
arcanobacterium	46	44	95.7%	0
arcobacter	26	21	80.8%	0

Genus	Num Samples	Num Correct	Species % Correct	Num False Pos.
arthrobacter	23	23	100.0%	0
bacillus	461	434	94.1%	1
bartonella	15	15	100.0%	0
bergeyella	10	10	100.0%	0
bordetella	98	82	83.7%	0
branhamella	25	24	96.0%	0
brevibacillus	23	21	91.3%	0
brevundimonas	48	46	95.8%	0
brucella	85	81	95.3%	0
burkholderia	117	106	90.6%	0
campylobacter	259	231	89.2%	1
cardiobacterium	8	7	87.5%	0
cdc group ef4	10	8	80.0%	0
cellulosimicrobium	10	10	100.0%	0
chromobacterium	6	5	83.3%	0
chryseobacterium	63	55	87.3%	3
chryseomonas	11	6	54.5%	0
citrobacter	46	31	67.4%	0
comamonas	36	20	55.6%	1
corynebacterium	465	409	88.0%	0
delftia	43	22	51.2%	2
dermabacter	13	13	100.0%	0
dermatophilus	12	12	100.0%	0
dysgonomonas	4	2	50.0%	0
edwardsiella	23	23	100.0%	0
eikenella	14	13	92.9%	0
empedobacter	31	30	96.8%	0
enterobacter	149	100	67.1%	0
enterococcus	325	309	95.1%	0
erysipelothrix	37	34	91.9%	0
escherichia	126	118	93.7%	0
flavimonas	19	11	57.9%	0
flavobacterium	6	5	83.3%	0
fluoribacter	39	25	64.1%	0
francisella	32	32	100.0%	0
gardnerella	28	28	100.0%	0
gemella	8	6	75.0%	0
gordonia	36	33	91.7%	0
grimontia	8	8	100.0%	0
haemophilus	275	188	68.4%	0
hafnia	15	7	46.7%	0
helicobacter	74	67	90.5%	0
jonesia	6	6	100.0%	0
kingella	31	30	96.8%	0
klebsiella	99	56	56.6%	1
kluyvera	16	11	68.8%	1
kocuria	61	59	96.7%	0
lactobacillus	74	74	100.0%	0
lechevalieria	16	16	100.0%	0
leclercia	7	5	71.4%	0
legionella	369	337	91.3%	0
leuconostoc	6	5	83.3%	0
listeria	118	93	78.8%	1
macrococcus	31	31	100.0%	0
mannheimia	12	7	58.3%	0
methylobacterium	59	41	69.5%	0
microbacterium	12	12	100.0%	0
micrococcus	77	74	96.1%	0

Genus	Num Samples	Num Correct	Species % Correct	Num False Pos.
midi calibration mix 1	25	25	100.0%	0
moraxella	151	140	92.7%	1
morganella	8	5	62.5%	0
mycobacterium	129	126	97.7%	0
myroides	33	33	100.0%	0
neisseria	172	115	66.9%	3
nocardia	123	109	88.6%	0
nocardioides	20	20	100.0%	0
nocardiopsis	12	12	100.0%	0
ochrobactrum	23	23	100.0%	0
oligella	19	9	47.4%	0
paenibacillus	28	28	100.0%	0
pandoraea	19	18	94.7%	0
pantoea	14	9	64.3%	0
pasteurella	87	59	67.8%	1
photobacterium	11	10	90.9%	0
plesiomonas	21	20	95.2%	0
proteus	66	45	68.2%	1
providencia	59	43	72.9%	0
pseudomonas	309	258	83.5%	6
psychrobacter	42	37	88.1%	0
rahnella	20	16	80.0%	0
ralstonia	66	57	86.4%	0
raoultella	9	7	77.8%	0
rhizobium	22	15	68.2%	0
rhodococcus	95	82	86.3%	0
riemerella	19	19	100.0%	0
roseomonas	50	44	88.0%	1
rothia	25	25	100.0%	0
saccharothrix	24	24	100.0%	0
salmonella	41	31	75.6%	0
serratia	65	34	52.3%	0
shewanella	37	36	97.3%	0
shigella	65	45	69.2%	0
sphingobacterium	66	46	69.7%	1
sphingomonas	20	12	60.0%	0
staphylococcus	1970	1696	86.1%	0
stenotrophomonas	74	74	100.0%	0
streptococcus	630	548	87.0%	0
suttonella	16	15	93.8%	0
tatlockia	42	38	90.5%	0
tsukamurella	10	9	90.0%	0
vibrio	97	89	91.8%	0
virgibacillus	7	7	100.0%	0
weeksella	17	17	100.0%	0
yersinia	136	131	96.3%	0
Average over 116			84.9%	
NOOHC6 Total	9706	8285	85.4%	27

MIDI is continuing the process of evaluation through replicated analysis of hundreds of species of bacteria using the FAST protocol and then checking the FAST data against the current database. The data collected using the FAST protocol is being used to create a FAST library (Table 5). The performance of the library constructed from FAST data will be substantially better as it will exactly match the fatty acid recovery profiles of the FAST extraction procedure. MIDI plans to continue this work without additional funding.

Table 5: Listing of Species in the Preliminary FAST Library

#	Name	#	Name
1	A	91	Neisseria-flavescens
2	Achromobacter-xylosoxidans-denitrificans	92	Neisseria-subflava
3	Achromobacter-xylosoxydans-xylosoxydans	93	Neisseria-weaverii
4	Acidovorax-delafieldii	94	Nocardia-asteroides
5	Acinetobacter-baumanii	95	Nocardia-brasiliensis
6	Acinetobacter-haemo	96	Nocardia-brevicatena
7	Acinetobacter-johnsonii	97	Nocardia-farcinica
8	Acinetobacter-lwoffii	98	Nocardia-otitiditis
9	Actinobacillus-equuli	99	Nocardia-pseud
10	Actinobacillus-ligni	100	Nocardia-transvaalensis
11	Actinobacillus-seminis	101	Nocardioides-albus
12	Actinobacillus-urea	102	Nocardiopsis-dassonvillei(wrong??)
13	Aeromonas-caviae	103	Ochrabactrum-anthropi
14	Aeromonas-hydrophila-hydrophila	104	Oligella-ureolytica
15	Aeromonas-veronii	105	Oligella-urethralis
16	Alcaligenes-faecalis	106	Paenibacillus-macerans
17	Amycolatopsis-orientalis	107	Paenibacillus-polymyxa
18	Arcabacterium-pyogenes	108	Pantoea-agglomerans
19	Arcanobacterium-haemolyticum	109	Pantoeae-agglommerans
20	Bacillus cereus	110	Pasteurella-canis
21	Bacillus-atropheus	111	Pasteurella-multocida-multocida
22	Bacillus-circulans	112	Plesiomonas-shige
23	Bacillus-megaterium	113	Proteus-mirabilis
24	Bacillus-mycoides	114	Proteus-penne
25	Bacillus-pumilus	115	Providencia-alcalifaciens
26	Bacillus-subtilis	116	Providencia-rettgeri
27	Bacillus-thuringiensis	117	Providencia-rustigiana
28	Bordatella-parapertussis	118	Providencia-stuartii
29	Bordetella-avium	119	Pseudomonas-aeruginosa
30	Bordetella-bronchiseptica	120	Pseudomonas-alcaligenes
31	Brevibacillus-parabrevis	121	Pseudomonas-mendocino
32	Brevundimonas-brevis	122	Pseudomonas-pseudoalcaligenes
33	Burkholderia cepacia	123	Pseudomonas-putida
34	Burkholderia-gladiolii	124	Pseudomonas-stutzeri
35	Burkholderia-multivorans	125	Ralstonia-pickettii
36	Chryseobacterium-meningosepticum	126	Rhodococcus-equi
37	Chryseomonas-luteo	127	Rhodococcus-rhodnii
38	Chrysobacterium-indologenes	128	Rhodococcus-rhodochrous
39	Citrobacter-amalo	129	Riemerella-anatipestifer
40	Citrobacter-braak	130	Roseomonas-cervi
41	Citrobacter-farmeri	131	Roseomonas-gilardii
42	Citrobacter-freundii	132	Rothia-dentocariosa
43	Comamonas-terrigena	133	Salmonella-cholerasuis-cholerasuis
44	Comamonas-testero	134	Salmonella-enteritidis
45	Corynebacterium-coyleae	135	Salmonella-typhi

#	Name	#	Name
46	Corynebacterium-jeikeium	136	Serratia-fonticola
47	Corynebacterium-matru	137	Serratia-marcescens
48	Corynebacterium-pseudodiphtheriticum	138	Serratia-plymuthica
49	Corynebacterium-pseudotuberculosis	139	Shewanella-putrifaciens
50	Corynebacterium-renale	140	Shigella-boydii
51	Corynebacterium-striatum	141	Shigella-dysenteriae
52	Corynebacterium-xerosis	142	Shigella-sonnei
53	Cupriavidus	143	Sphingobacterium multivorans
54	Dermobacter-hominis	144	Sphingomonas-paucimobilis
55	Edwardsiella-tarda	145	Staphylococcus aureus
56	Eikenella-corrodens	146	Staphylococcus-auricularis
57	Enterobacter-aerogenes	147	Staphylococcus-capitis-capitis
58	Enterobacter-cloacae	148	Staphylococcus-chromogenes
59	Enterobacter-sakazaii	149	Staphylococcus-cohnii
60	Enterococcus-casseliflavus	150	Staphylococcus-epidermidis
61	Enterococcus-cecor	151	Staphylococcus-haemolyticus
62	Enterococcus-dispar	152	Staphylococcus-hominis-novobioticus
63	Enterococcus-durans	153	Staphylococcus-intermedius
64	Enterococcus-faecalis	154	Staphylococcus-lentus
65	Enterococcus-faecium	155	Staphylococcus-saprophyticus
66	Escherichia coli	156	Staphylococcus-sciuri-sciuri
67	Flavimonas-oryzihabitans	157	Staphylococcus-simulans
68	Gordonia-aichi	158	Staphylococcus-xylosus
69	Gordonia-sputi	159	Stenotrophomonas maltophilia
70	Kingella-denitrificans	160	Streptococcus-anginosus
71	Klebsiella pneumoniae pneumoniae	161	Streptococcus-canis
72	Klebsiella-oxytoca	162	Streptococcus-mitis
73	Kocuria-kristinae	163	Streptococcus-mutans
74	Kocuria-rosea	164	Streptococcus-parasuis
75	Kocuria-varians	165	Streptococcus-pneumoniae
76	Lactobacillus-delbrucki-bulgaricus	166	Streptococcus-porcini
77	Listeria-grayi	167	Streptococcus-salivarius
78	Listeria-innocua	168	Streptococcus-sanguis
79	Listeria-ivanovii-ivanovii	169	Streptococcus-uberis
80	Listeria-monocytogenes	170	Suttonella-indologenes
81	Listeria-seeligeri	171	Vibrio-parahaemolyticus
82	Macrococcus-caseolyticus	172	Vibrio-vulnificus
83	Micrococcus-luteus	173	Virgibacillus-pantothenticus
84	Micrococcus-lylae	174	Weeksella-virosa
85	MIDI Calibration Mix RNOOH6	175	Yersinia-enterocolitica-enterocolitica
86	Moraxella-bovis	176	Yersinia-frede
87	Moraxella-canis	177	Yersinia-inter
88	Morganella-morganii	178	Yersinia-kristin
89	Mycobacterium smegmatis	179	Yersinia-pseudotuberculosis
90	Neisseria-cinerea		

To make the library more comprehensive, MIDI acquired a total of 784 new bacterial strains from 539 species. The fatty acid profiles for these strains were analyzed and will be used to improve the Sherlock clinical library. Table 6 summarizes the strains by genus and species. Of the 539 species, 397 are new species for the library. The other 142 species will strengthen the existing entries in the library for those species. So far, 248 new entries have already been added to the existing library. MIDI will continue this work without additional funding.

It is important to be sure that the identities of strains added to a library are correct. To confirm the identities of these strains, we had 16s rRNA sequencing performed. The sequence data provides a second method of identification.

Table 6: New Bacterial Strains Acquired and Analyzed for Addition to Clinical Library

Genus	#strains	#species	Species
Abiotrophia	1	1	defectiva
Achromobacter	6	2	piechaudii xylosoxidans
Acidithiobacillius	1	1	ferrooxidans
Acidovorax	3	3	defluvii delafieldii valerianellae
Acinetobacter	28	21	baumannii calcoaceticus genomospecies 10 genomospecies 11 genomospecies 14 genomospecies 16 genomospecies 3 genomospecies 6 genomospecies 9 haemolyticus johnsonii junii lwoffii radioresistens sp. baylyi parvus schindleri ursingii grimontii towneri
Actinobacillus	1	1	dephinicola
Actinocorallia	3	2	aurantiaca glomerata
Actinomyces	1	1	naeslundii
Actinoplanes	1	1	regularis
Aerococcus	1	1	urinae
Aeromicrobium	1	1	erythreum
Aeromonas	6	6	caviae hydrophila molluscorum simiae sobria veronii
Afipia	2	2	birgiae clevelandensis
Agrococcus	1	1	citreus
Agromyces	3	3	cerinus humatus rhizospherae
Alcaligenes	2	2	faecalis xylosoxydans
Alcanivorax	1	1	borkumensis
Algoriphagus	1	1	halophilus
Alicaligenes	1	1	defragrans
Alicyclobacillus	4	4	herbarius pomorum tolerans vulcanalis
Aminobacter	2	2	aganoensis aminovorans
Amycolatopsis	1	1	orientalis
Aneurinibacillus	1	1	thermoaerophilus
Arcobacter	1	1	cibarius
Arthrobacter	3	3	cumminsii russicus stackebrandtii
Aureobacterium	1	1	barkeri
Azotobacter	1	1	vinelandii
Bacillus	33	29	acidicola arenosi arvi barbaricus cibi circulans decolorationis endophyticus farraginis fordii funiculus galactosidilyticus horti humi muralis nealsonii neidei okuhidensis patagoniensis pseudomycoides psychrodurans psychrotolerans pycnus saliphilus silvestris siralis sonorensis subterraneus weihenstephanenesis
Blastomonas	1	1	natatoria
Bordetella	5	4	avium bronchiseptica hinzii parapertussis
Borrelia	3	3	burgdorferi garinii turcica
Bosea	1	1	minatitlanensis
Brachybacterium	1	1	muris
Brevibacillus	6	6	brevis formosus invocatus limnophilus parabrevis thermoruber
Brevibacterium	11	9	epidermis casei iodinum linens luteolum lyticum mcbrellneri paucivorans sanguinis
Brevundimonas	2	2	bacteroides nasdae

Genus	#strains	#species	Species
Burkholderia	53	26	ambifaria andropogonis anthina caledonica caryophylli cenocepacia cepacia dolosa fungorum gladioli glathei graminis kururiensis multivorans phenazinium phymatum phytofirmans pyrrocinia sacchari stabilis thailandensis tropica ubonensis unamae vietnamiensis xenovorans
Campylobacter	1	1	lari
Capnocytophaga	2	2	canimorsus cynodegmi
Cellulomonas	6	4	biazotea cartae cellasea fimi
Cellulophaga	1	1	lytica
Chryseobacterium	5	5	balustinum defluvii gleum indologenes meningosepticum
Chryseomonas	2	1	luteola
Citrobacter	1	1	murliniae
Comamonas	3	3	kerstersii terrigena testosteroni
Corynebacterium	7	6	auriscanis durum felinum macginleyi propinqum singulare
Cupriavidus	1	1	necator
Deinococcus	1	1	murrayi
Dermabacter	1	1	hominis
Dermatophilus	1	1	chelonae
Edwardsiella	1	1	tarda
Elizabethkingia	1	1	miricola
Enteric	1	1	CDC_group
Enterobacter	2	2	hormaechei kobei
Enterococcous	2	2	ratti villorum
Enterococcus	56	20	asini avium canis casseliflavus cecorum columbae dispar durans faecalis faecium flacescens gallinarum hermanniensis hirae malodoratus mundtii pseudoavium raffinosus saccharolyticus sulfureus
Escherichia	12	4	blattae coli fergusonii vulneris
Exiguobacterium	1	1	undae
Flavimonas	1	1	oryzihabitans
Flavobacterium	5	5	acidurans branchiophilum columnare mizutaii pectinovorum
Gemella	1	1	cuniculi
Geobacillus	8	8	caldoxyosilyticus debilis gargensis tepidamans thermodenitrificans toebii uzenensis vulcani
Gordonia	1	1	otitidis
Gracilibacillus	1	1	dipsosauri
Haemophilus	8	3	actinomycetemcomitans aphrophilus pittmaniae
Halomonas	2	2	alimentaria aquamarina
Helcococcus	3	2	kunzii ovis
Hydrogenophaga	2	2	atypica palleronii
Hylemonella	1	1	gracilis
Janibacter	1	1	terrae
Klebsiella	20	5	oxytoca planticola pneumoniae singaporensis terrigena
Kluyvera	3	1	cryocrescens
Kocuria	7	5	carniphila kristinae rhizophila rosea varians
Kribbella	1	1	solani
Kytococcus	1	1	schroeteri
Lactobacillus	47	29	acidophilus agilis alimentarius amylovorus animalis bifermentans brevis buchneri casei casei delbrueckii fructivorans gasseri helveticus homohiochii intestinalis jensenii johnsonii lactis murinus parabuchneri paracasei plantarum reuteri rhamnosus ruminis salivarius sharpeae vaginalis
Lactococcus	1	1	lactis
Lautropia	1	1	mirabilis
Lentzea	1	1	californiensis
Leptospirillum	1	1	ferriphilum
Leuconostoc	11	5	carnosum citreum lactis mesenteroides pseudomesenteroides
Listeria	18	6	grayi innocua ivanovii monocytogenes seeligeri welshimeri
Luteococcus	1	1	sanguinis

Genus	#strains	#species	Species
Marinomonas	1	1	vaga
Methylobacterium	1	1	aquaticum
Methylosarcina	1	1	fibrata
Microbacterium	26	26	arborescens aurantiacum aurum chocolatum dextranolyticum esteraromaticum flavescens imperiale keratanolyticum lacticum laevaniformans liquefaciens luteolum maritypicum oxydans saperdae schleiferi sp. testaceum trichothecenolyticum aerolatum hydrocarbonoxydans paraoxydans phyllosphaerae resistens ulmi
Micrococcus	16	8	aurantiacus candicans conglomeratus diversus freudenreichii luteus lylae naucinus
Moraxella	4	4	boevrei bovis canis oblonga
Morganella	1	1	morganii
Mycobacterium	20	20	boenickei brisbanense canariasense chimaera cosmeticum doricum florentinum gordonae holsaticum houstonense immunogenum lacus nebraskense neworleansense palustre parascrofulaceum parmense saskatchewanense szulgai ulcerans
Mycoplana	1	1	ramosa
Neisseria	9	6	cinerea elongata flavescens iguanae subflava weaveri
Nocardia	36	26	abscessus africana anaemiae aobensis araoensis asteroides arthritidis asiatica concava crassostreae cyriacigeorgica elegans fluminea higoensis ignorata inohanensis kruczakiae niigatensis otitidiscaviarum paucivorans pneumniae puris thailandica vermiculata veterana yamanashiensis
Nocardioides	2	2	jensenii plantarum
Nocardiopsis	3	3	exhalans prasina trehalosi
Oceanobacillus	1	1	picturae
Oceanospirillum	1	1	maris
Ochrobactrum	1	1	gallinifaecis
Ornithobacterium	1	1	rhinotracheale
Paenibacillus	26	26	agarexedens amylolyticus anaericanus barcinonensis borealis brasilensis chibensis chinjuensis cineris cookii favisporus glycanilyticus graminis illinoisensis kobensis lactis massiliensis naphthalenovarns nematophilus odorifer phyllosphaerae rhizosphaerae stellifer terrae timonensis turicensis
Pandoraea	3	3	apista norimbergensis pnomenusa
Pantoea	1	1	citrea
Pasteurella	4	4	canis dagmatis multocida volantium
Pediococcus	10	5	acidilactici damnosus dextrinicus parvulus pentosaceus
Proteus	2	2	mirabilis penneri
Providencia	4	4	alcalifaciens rettgeri rustigianii stuartii
Pseudomonas	42	33	alcaligenes avellanae beijerinckii brassicacearum brenneri cedrina congelans constantinii cremoricolorata extremorientalis fluorescens frederiksbergensis gessardii graminis grimontii indica kilonensis lini lutea mendocina monteilii mosselii orientalis palleroniana poae proteolytica pseudoalcaligenes putida rhizosphaerae stutzeri thermotolerans thivervalensis trivialis
Pseudonocardia	1	1	autotrophica
Ralstonia	2	2	eutropha pickettii
Rhodococcus	12	5	coprophilus gordoniae rhodochrous ruber triatomae
Riemerella	1	1	anatipestifer
Roseomonas	2	2	cervicalis gilardii
Rothia	1	1	dentocariosa
Rubrobacter	1	1	radiotolerans
Salmonella	1	1	choleraesuis
Serratia	22	10	entomophila ficaria fonticola grimesii marcescens odorifera plymuthica proteamaculans rubidaea rubidea
Shigella	10	3	boydii dysenteriae sonnei
Staphylococcus	5	2	caprae epidermidis
Streptococcus	50	25	agalactiae alactolyticus anginosus bovis canis constellatus dysgalactiae

Genus	#strains	#species	Species
			equi equinus equisimilis hyointestinalis hyovaginalis infantis intermedius mitis mutans oralis parasanguinis peroris pneumoniae pyogenes salivarius sanguinis sobrinus uberis
Streptomyces	1	1	purpureus
Taylorella	1	1	equigenitalis
Telluria	2	1	mixta
Vibrio	1	1	harveyi
Virgibacillus	1	1	proomii
Yersinia	3	3	aleksiciae enterocolitica frederiksenii
Yokenella	1	1	regensburgei
Total	**784**	**539**	

Objective 3

Enhance Sherlock from a standalone application with a proprietary data store to a secure client-server application with an underlying relational database, allowing Sherlock BioTerNet to rapidly track bacterial strains across multiple sites in real time.

Software work toward this objective involved three aspects: evaluation of new technologies necessary to meet the objective; definition of a database to maintain the necessary information; and development of a prototype system that allows access for multiple users at remote sites using Internet capabilities. Simultaneously, improvements to the base Sherlock system were implemented to accommodate these needs.

Evaluation of new technologies

- **Database Tools**
  Available database tools were evaluated for suitability with the goals of Sherlock BioTerNet. SQL Server was selected as the target database tool for Sherlock BioTerNet. (The lightweight MSDE version of SQL Server can be used for individual Sherlock stations; the full SQL Server can be used for the shared BioTerNet database server.) A SQL Server database system was established at MIDI Inc. for exploration of techniques. The prototype was developed using this database tool.

- **Web-based access**
  User Interface tools for web-based access to the Sherlock BioTerNet system were evaluated. The objective was to select a set of standard user interface tools. Microsoft's Visual Studio 2005 (with .Net framework and the C# computer language) was selected as the development tool of choice, given its rich set of built-in components. The prototype was constructed within this development environment.

- **XML import/export techniques**
  An evaluation of XML for serialization of Sherlock objects concluded that this technique could be effectively used. Individual sample results could easily be exported and imported using XML, based on built-in capabilities of Windows .Net. This technique would have been explored further in years two and three of the project but was not used in year one.

- **Alternative instrument acquisition software**
  An alternative instrument acquisition solution, using EZChrom software from Agilent Technologies, was evaluated. This software is more flexible than the current ChemStation software, allowing new instruments to be supported. The initial evaluation demonstrated that results equivalent to current ChemStation results can be achieved using EZChrom. . This technique would have been explored further in years two and three of the project but was not used in year one

- **Access to up-to-date taxonomic information**
  Working with Professor George Garrity of Michigan State University, the BioTerNet prototype has been

linked to a "Names For Life" ™ server, an online resource that gives taxonomic information associated with the identification received from BioTerNet. This technique assures that BioTerNet will maintain up-to-date taxonomic information automatically.

Database Definition

- **Requirements**
  Detailed requirements for the database definition were developed.
  These requirements recognize the need to redefine first-level objects of the Sherlock system in order to meet the needs of Sherlock BioTerNet.

- **Design**
  Redesign of Sherlock first-level objects is necessary to fulfill the more flexible needs of Sherlock BioTerNet. The older Sherlock "Method" object is to be replaced with "Protocol / Transform / Test" objects which will handle a broader variety of experiments performed upon Sherlock samples. The older Sherlock "Profile" object is to be replaced with the "Results" object which will store a wide set of result formats. The older Sherlock "Library / Library Element" objects are to be replaced by "Organism / Result" objects which allow polyphasic information to be stored for each bacterial species. The net effect will be a much more flexible data definition.

- **Schema Design**
  A database schema was developed to allow the requirement and design elements to be embodied in a SQL server database.

System Prototype

A prototype has been developed that demonstrates Web access to a database encompassing the new Sherlock object design. A MSDE SQL Server database running the initial Sherlock BioTerNet schema has been developed. Further, this database has been scripted so that changes can be made and automatically disseminated to all database componentry.

- Database Componentry: software has been developed that allows access to the Sherlock BioTerNet database using a Microsoft .Net C# class library.
- Data Import: allows the prototype to import data from existing Sherlock system, and incorporate that data into the database, using a Web-based interface.
- Website: A BioTerNet website has been developed that allows accessing data in the BioTerNet database in report format, as well ass administration of the BioTerNet database. This website is a key result for the software project.
- DNA Extension: The Website allows polyphasic results to be viewed for samples that have both FAME and DNA data associated.
- Tracker: A tracking capability has been included in the Website to allow tracking a sample against all samples in the database, allowing determination of patterns within the data. Tracker is the first step toward tools that recognize unusual patterns within the Sherlock data.

Sherlock 6.0.

Several aspects of the work for BioTerNet are already providing value through updates to the existing Sherlock product base. Sherlock 6.0 was released in January 2006 with the following new capabilites:

- Sherlock 6.0 is the first polyphasic version of Sherlock, allowing cross-library reporting of DNA and FAME.

- Sherlock 6.0 includes the updated BTR3 and RBTR3 libraries, with a new *Bacillus-anthracis* subgroup and more near neighbors. Library updates include dozens of improved entries, as well as updated taxonomies.
- Sherlock 6.0 includes a *Cluster* capability for automatically identifying related samples.
- A "Framer" application has been developed allowing DNA data to be placed in the correct context for incorporation with Sherlock DNA.

Updates to Sherlock 6.0 have been made, allowing use of the new ChemStation B.01 and B.02 software with Sherlock. Expected release date of that product is October 2006.

FDA 510(k) Clearance

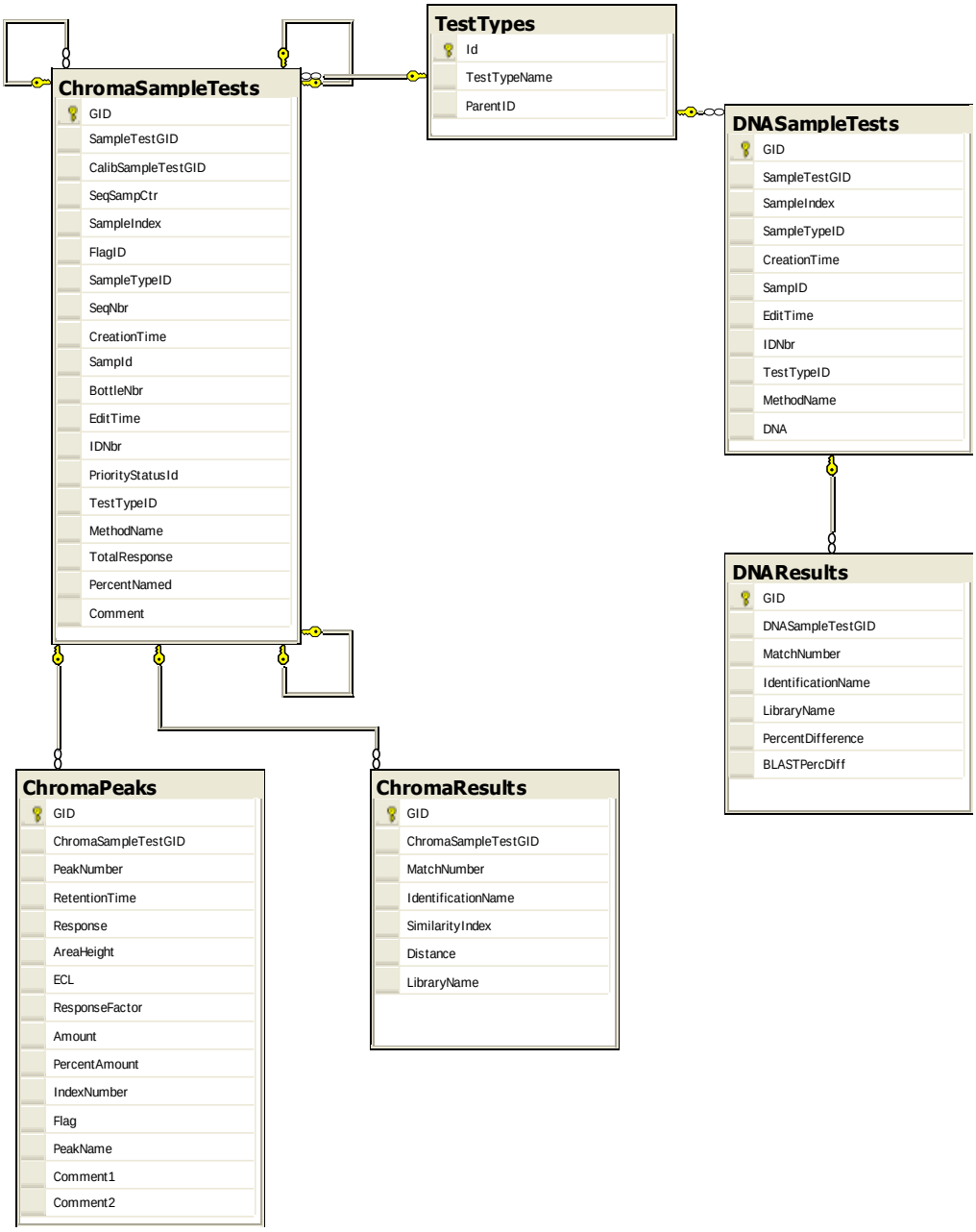
MIDI received FDA 510(k) clearance for the identification of anthrax based on the currently available extraction technique and method.

Database Design

The Sherlock BioTerNet database is based on a SQL server platform. A schema written for the platform gives access to the data from multiple PCs on a network. While the full schema contains over thirty tables, the key section of the schema is detailed in Figure 2.

The key section of the Sherlock BioTerNet database consists of *Tests* that hold the results for individual *Samples*. As shown in the figure XXX, there are two main types of *Tests*: *ChromaSampleTests* contain the results from chromatographic (FAME) tests; *DNASampleTests* contain the results from DNA (16S) tests. (Sherlock BioTerNet is designed for expandability as other types of tests could be added.) Details under each test type are unique to the needs of that test. For instance, the *ChromaSampleTests* table contains *BottleNumber* which relates to the bottle that was run on the Gas Chromatograph; the subsidiary table *ChromaPeaks* contains the details of individual peaks including *PeakName* (the name of the Fatty Acid) and *PercentAmount* (the percentage of that fatty acid). Finally, the *ChromaResults* table contains the identification against the library of known organisms with *IdentificationName* and *SimilarityIndex* being the key results.

Figure 2: Key Schema Tables for BioTerNet



In order to make the database usable from the Visual Studio .Net framework, two techniques are employed: Stored Procedures (SPROC) and a Data Access Layer (DAL). Stored Procedures are access functions that reside within the database; each table has a common set of access functions. The Data Access Layer resides within the .Net framework, exposing the tables in .Net and communicating to the database through SPROC. With SPROC and DAL, the database appears like a set of objects within the .Net framework, allowing C# application code to be written conveniently.

The source example below shows how easy it is to develop code once SPROC and DAL are available.

SPROC and DAL example

In this example, the application code wants to add a new sample to the system.
The application code writes something like this:

```
       DB_Samples Samp = new DB_Samples();
        Samp.Creator = txtCreator.Text;
        Samp.DateReceived= Convert.ToDateTime(txtDateReceived.Text);
        Samp.Description = txtDescription.Text;
        Samp.Save();
```

The "DB_Samples" object is an automatically generated DAL object looking like this:

```
    public class DB_Samples : DB
    {
…
        protected DateTime _DateReceived;
        public DateTime DateReceived
        {
            get { return _DateReceived; }
            set { _DateReceived = value; }
        }
…
        public void Save()
        {
…
   ExecuteNonQuery( StoredProcedure, "p_Samples_Save", params);
…
```

The StoredProcedure p_Samples_Save is an automatically generated SPROC. This procedure is stored directly in the database and looks like:

```
ALTER PROCEDURE [dbo].[p_Samples_Save]
…         @Creator varchar(100),
          @DateReceived smalldatetime,
          @Description varchar(7000)
…
                        UPDATE Samples SET
                          ,Creator = @Creator
                          ,DateReceived = @DateReceived
                          ,Description = @Description
```

This fully automates the process of creating new samples.

Without automation, one would need to update the SPROC and DAL code each time a table was changed in the database. To avoid this support nightmare, the third-party tool *CodeSmith* has been used. *CodeSmith* allows defining a set of templates for both SPROC and DAL. Using the information in the database, both SPROC and DAL can be automatically generated by *CodeSmith*, assuring that SPROC and DAL match the current database schema.

As a final automation tool, deployment scripts were developed. Deployment scripts allow the database to be developed on one PC and then deployed onto another. One simply runs the scripts on the target PC and the base BioTerNet database is constructed, ready for use.

Objective 4

Develop algorithms that automatically monitor BioTerNet's centralized database for suspicious events that may be indicative of a bioterrorism/biowarfare event.

No significant work was done on this objective as it was schedule for the second year of the project.

Prototype Use Case

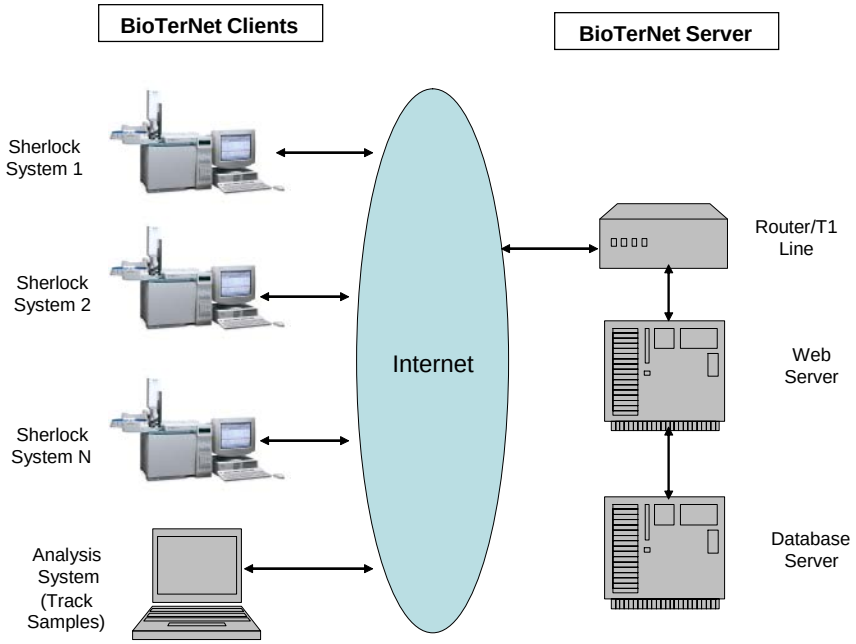
The following is a *use case* demonstrating the capabilities of the BioTerNet prototype. First a high level diagram showing how BioTerNet is organized is presented. This is followed by a detailed description showing how samples are loaded into the system, how sample information is displayed at a client computer and finally how samples are tracked using a client computer.

In its final form BioTerNet would consist of a client server system on the Internet. This will allow any user of a Sherlock System with an Internet connection to upload data to a centralized database through a Web Server. Also any person with the appropriate credentials (user name / password) will be able to track any sample across multiple laboratories in a matter of seconds.

The prototype of BioTerNet is implemented as a three-tier client server system. The first tier is a separate Database implemented using Microsoft SQL Server. All sample information is stored in the Database. The second tier is a web server implemented using the Microsoft Internet Information Server (IIS6). Clients only connect to the web server and only the web server connects to the database. This prevents users from having any direct contact with the data stored in the database.

For prototyping purposes, the web server is connected to MIDI's router/T1 line. This router/T1 line is attached to MIDI's Internet Service Provider allowing anyone with an Internet connection to access the web site. Both of these tiers are hosted on computers running Windows Server 2003. The last tier consists of any computer with an Internet connection. The prototype has been tested from MIDI's local area network and from computers that are not at MIDI using whatever Internet connection is available to the outside world. (Tests were made using computers connected to Comcast).

Figure 3: BioTerNet Prototype Architecture



Use Case Overview

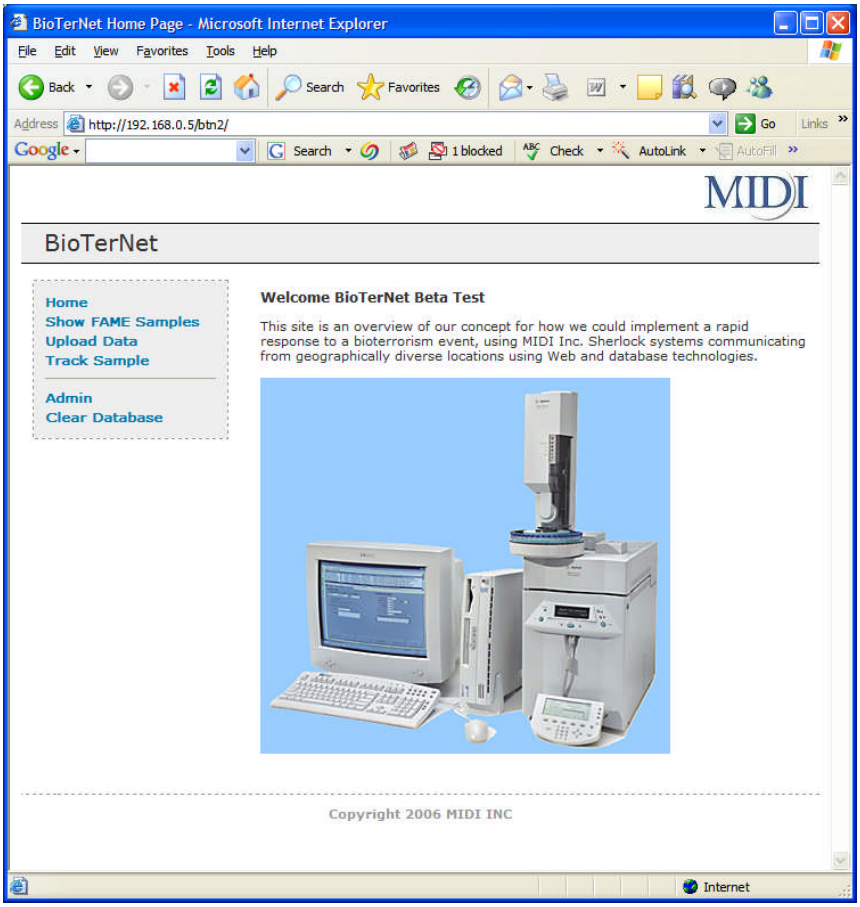
The use case demonstrates how to:

1 – Load sample information into BioTerNet from any Sherlock System that has access to the World Wide Web. (Please note that this can also be accomplished by moving a file from a secure system running the Sherlock System to any system that has internet access).

2 – View sample information in the BioTerNet database from a client computer.

3 – Track samples from a client computer. The data used in this example consists of anthrax data received from USAMRIID and anthrax data received from the Connecticut Department of Health. This example shows the BioTerNet prototype determining that the case of anthrax that a 94 year old Connecticut woman had in 2001 was the Ames strain of anthrax. The time to make this determination once the data was loaded into the prototype was well under one minute.

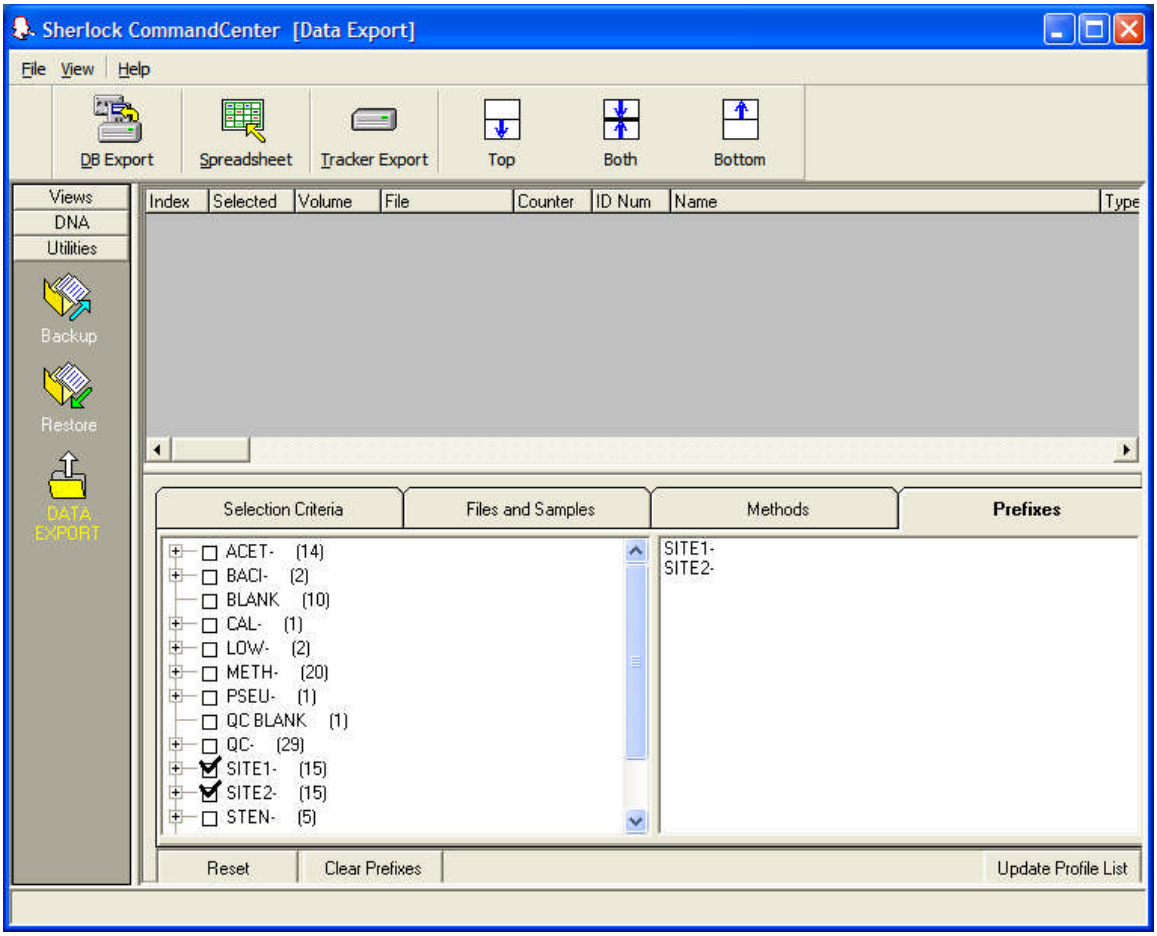
Top Level Screen - All operations are done using the BioTerNet Web site. The left hand side of the web site points the user to the various operations that the user can perform. The top level page of the website follows:



How to load sample information into the BioTerNet Centralized Database

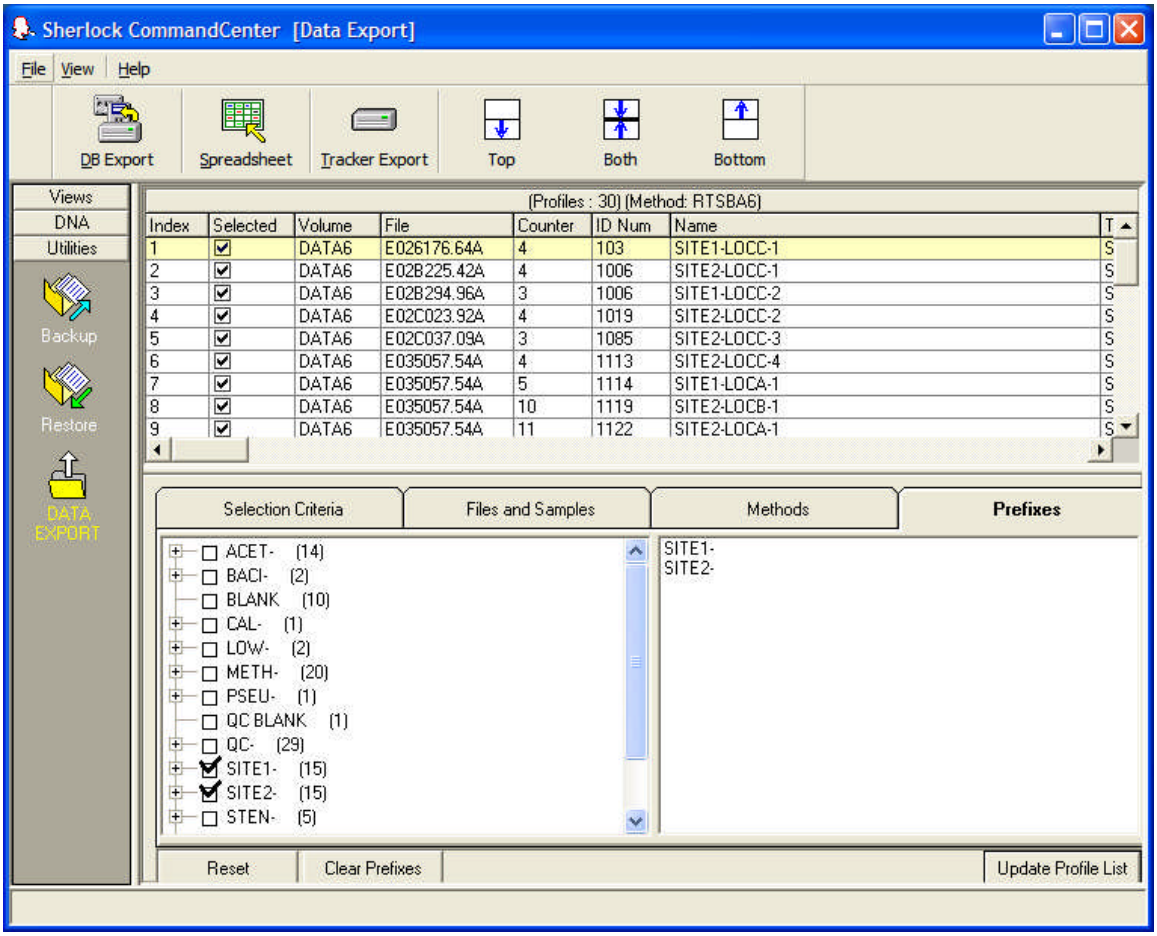
First samples are selected from a computer running Sherlock 6.0.

On a Sherlock System use, the Data Export applet to select samples that will be loaded into the centralized Database.



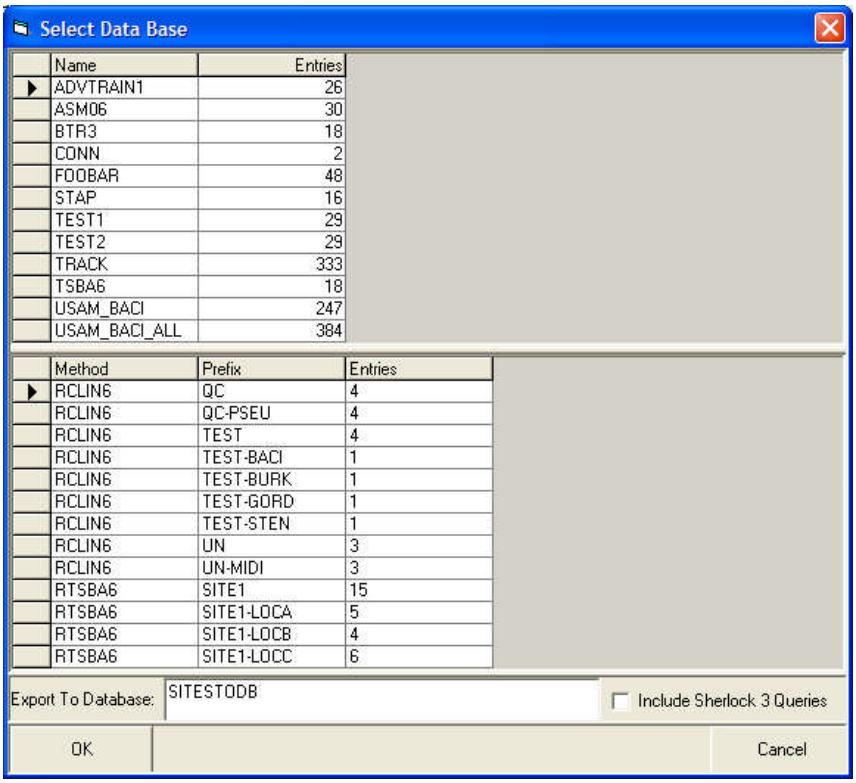
Profile selected samples

The profile operation is used to name/quantify the chemical compounds in the sample and to identify the samples using the Sherlock software.



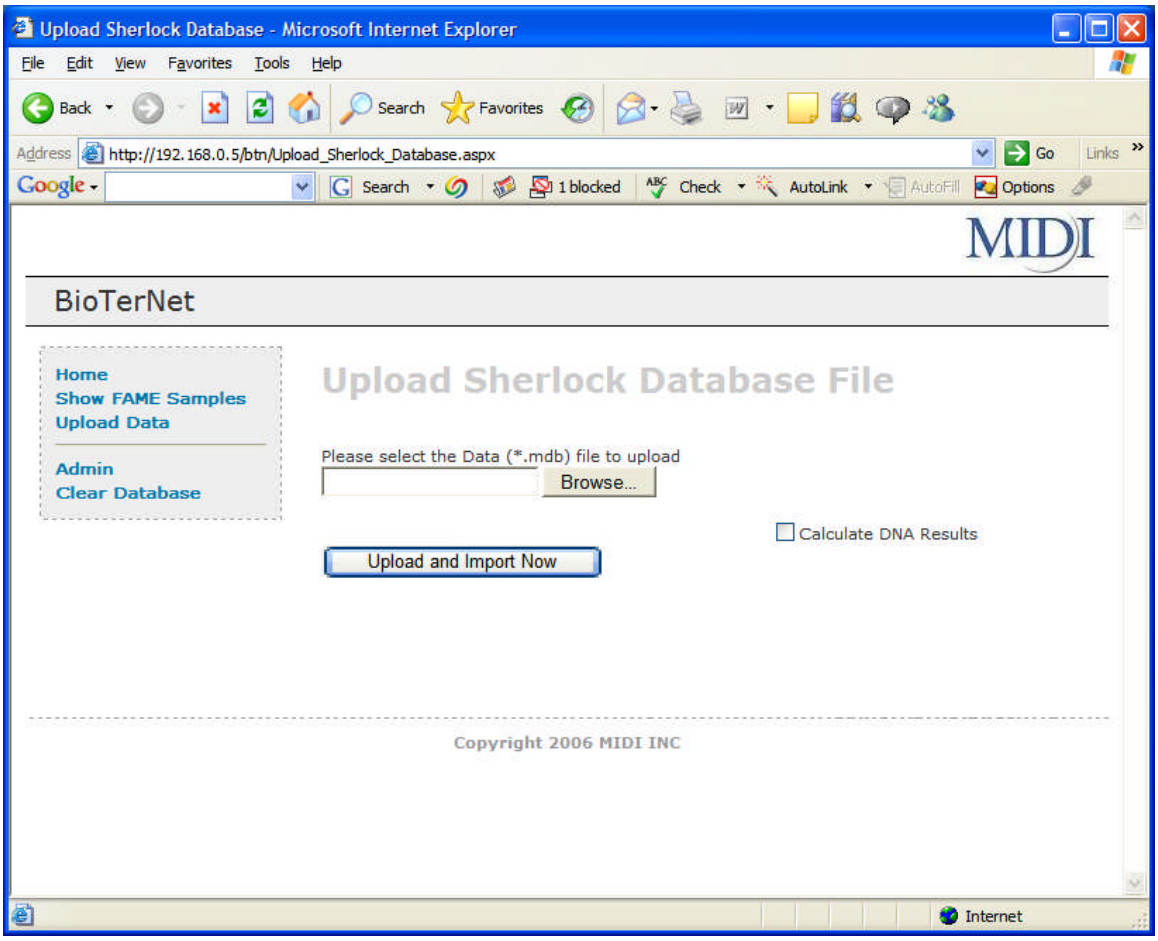
Create file to upload to the server through Sherlock's Data Export

The following screen shot shows the user creating the database file SITESTODB. This database file contains all the information for the profiled samples. Uploading this file to the BioTerNet database transfers the information into BioTerNet.

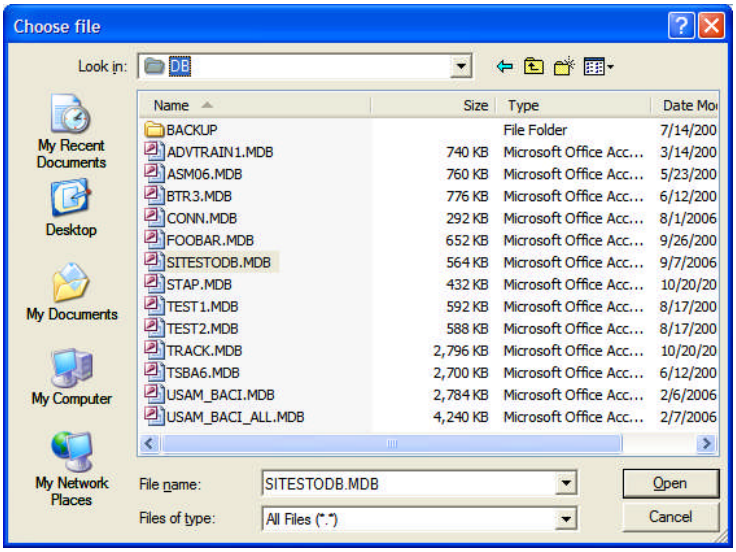


Uploading - Transferring data to BioTerNet

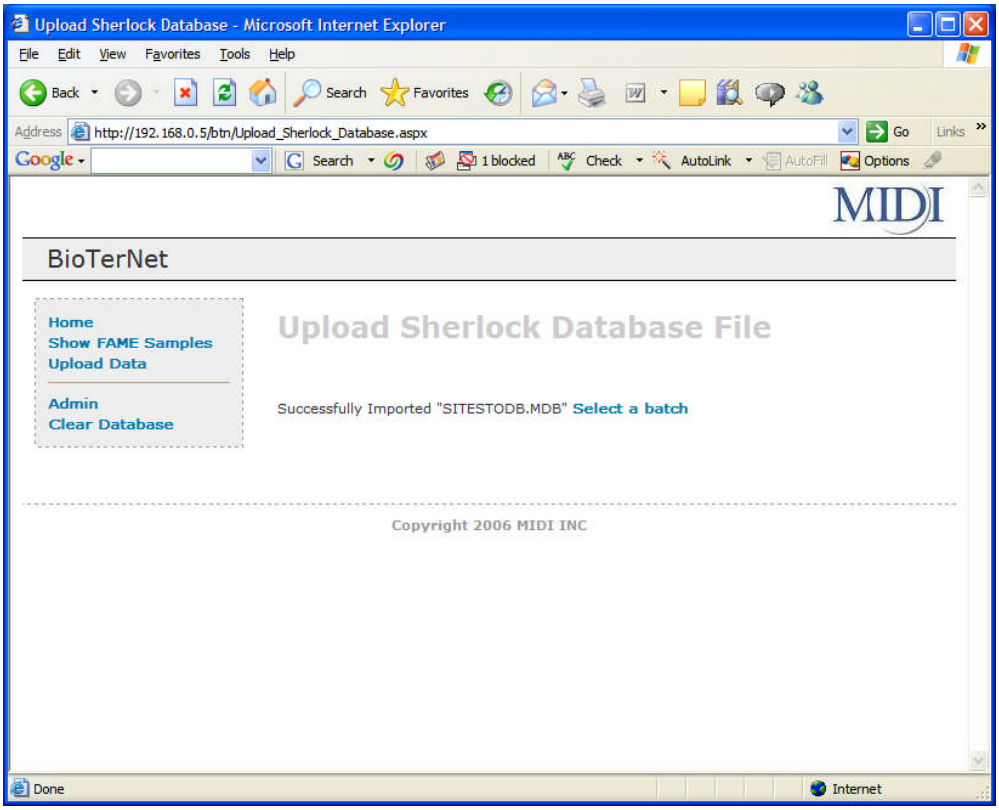
The information stored in the file SITESDB is loaded into the BioTerNet database from any computer that has an internet connection. Please note that it is not necessary for the computer that runs samples through Sherlock to be attached to the internet. The file SITESTODB could have been offloaded from the computer running Sherlock to another computer that has an internet connection.



Pressing the browse button, displays the following screen that allows the user to select the SITESTODB file.

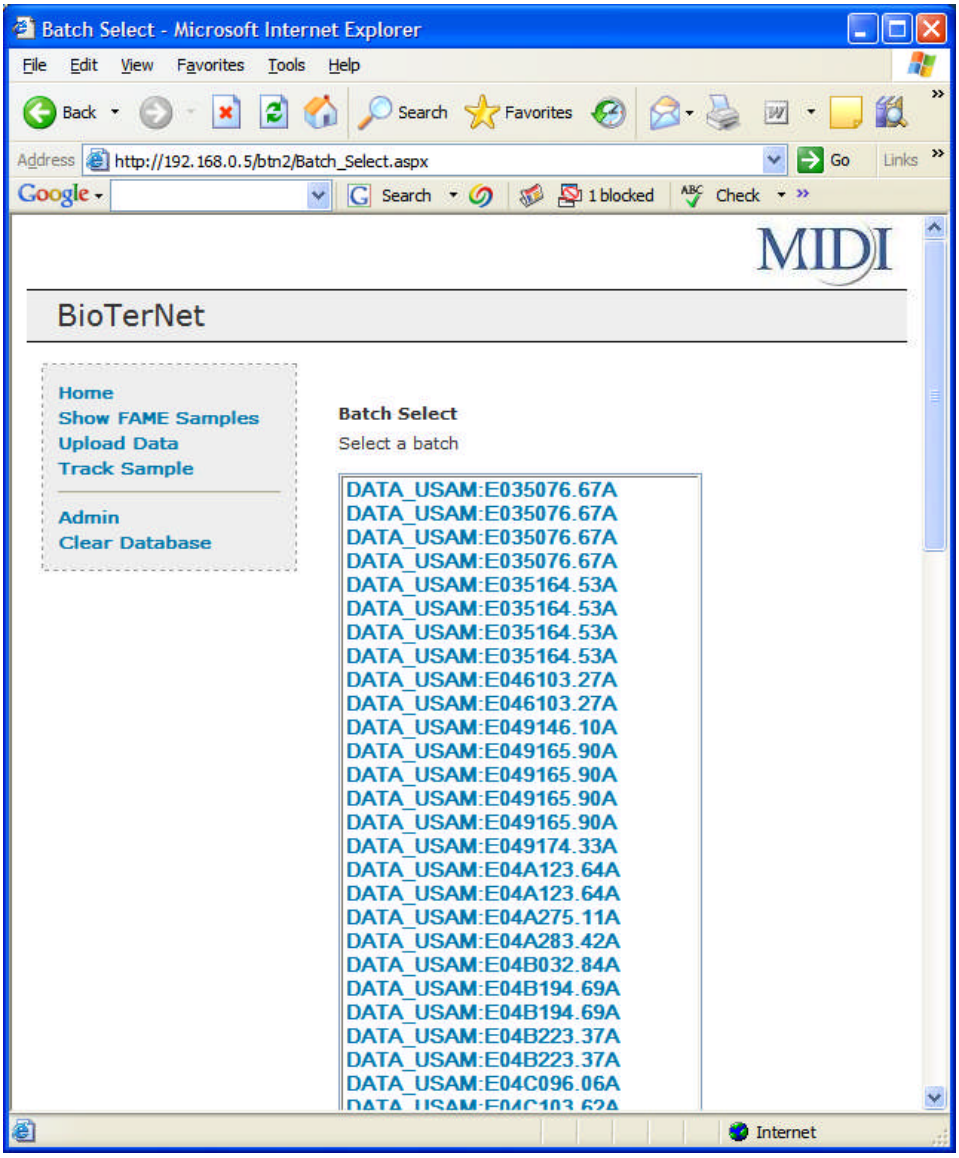


The following screen appears after the data has been uploaded to the BioTerNet database.



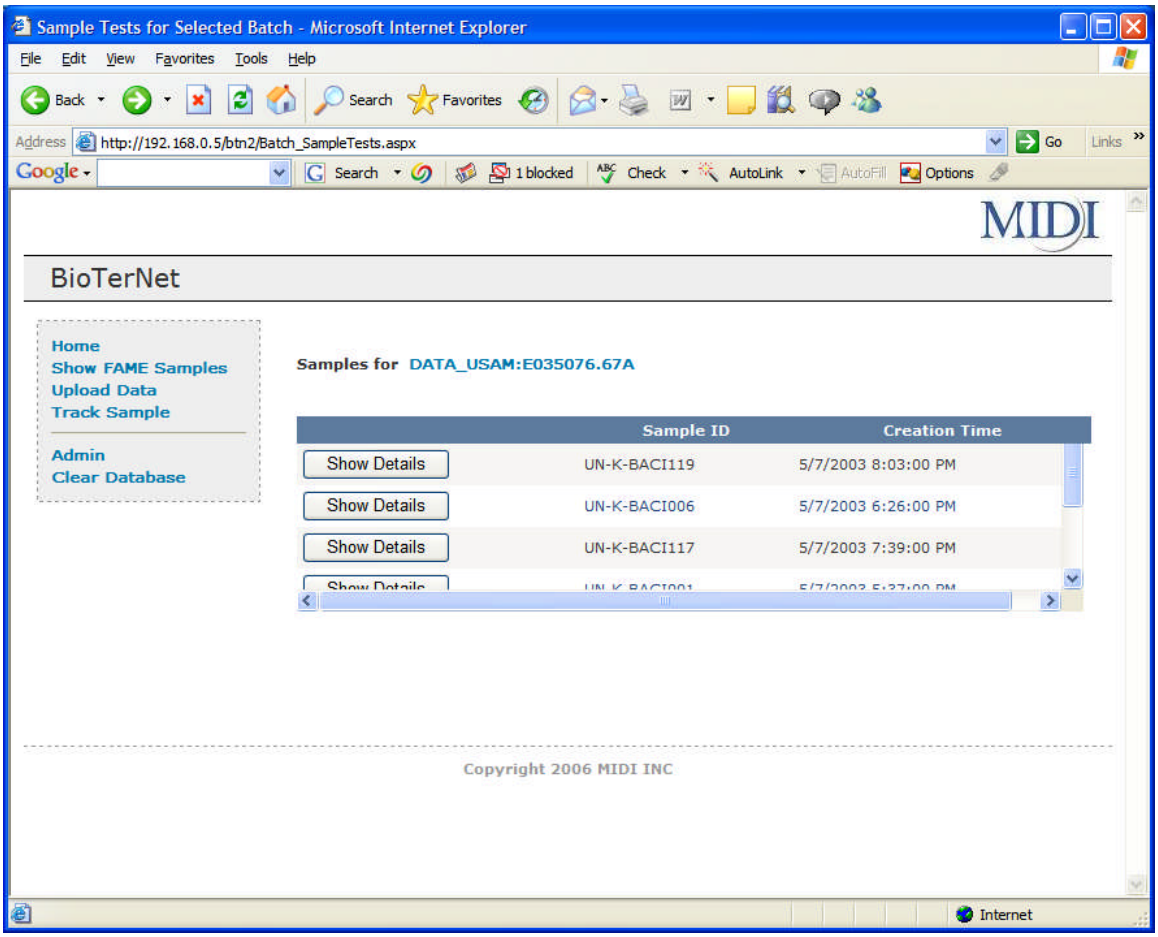
Display Sample information from the BioTerNet Database

Select a batch from the BioTerNet database via the BioTerNet web site. A batch consists of a series of samples that were run on the same instrument consecutively.



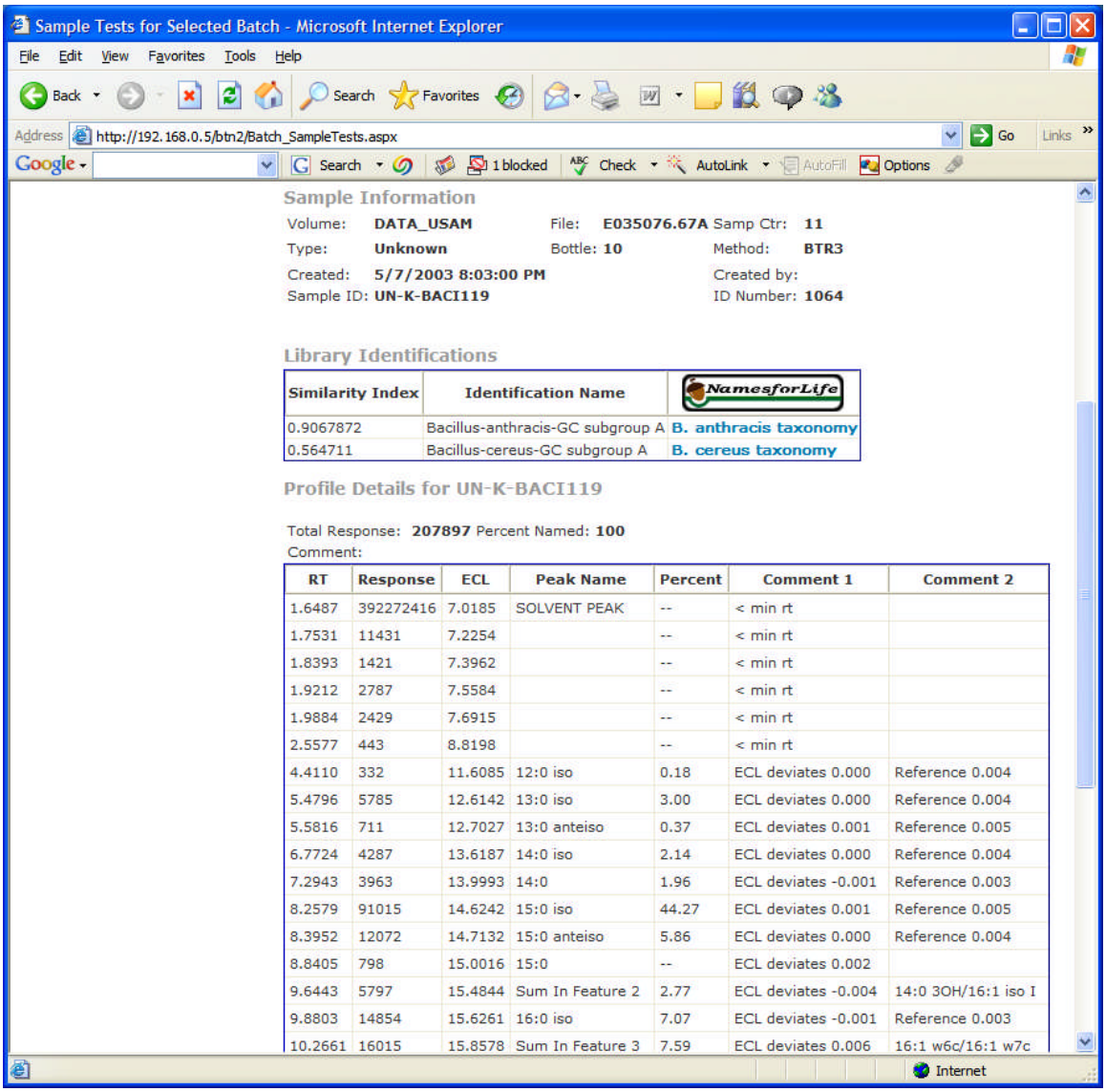
Select an individual sample

Samples are shown by sample ID and the time they were run. Press the Show Details button to get specific information for the desired sample.



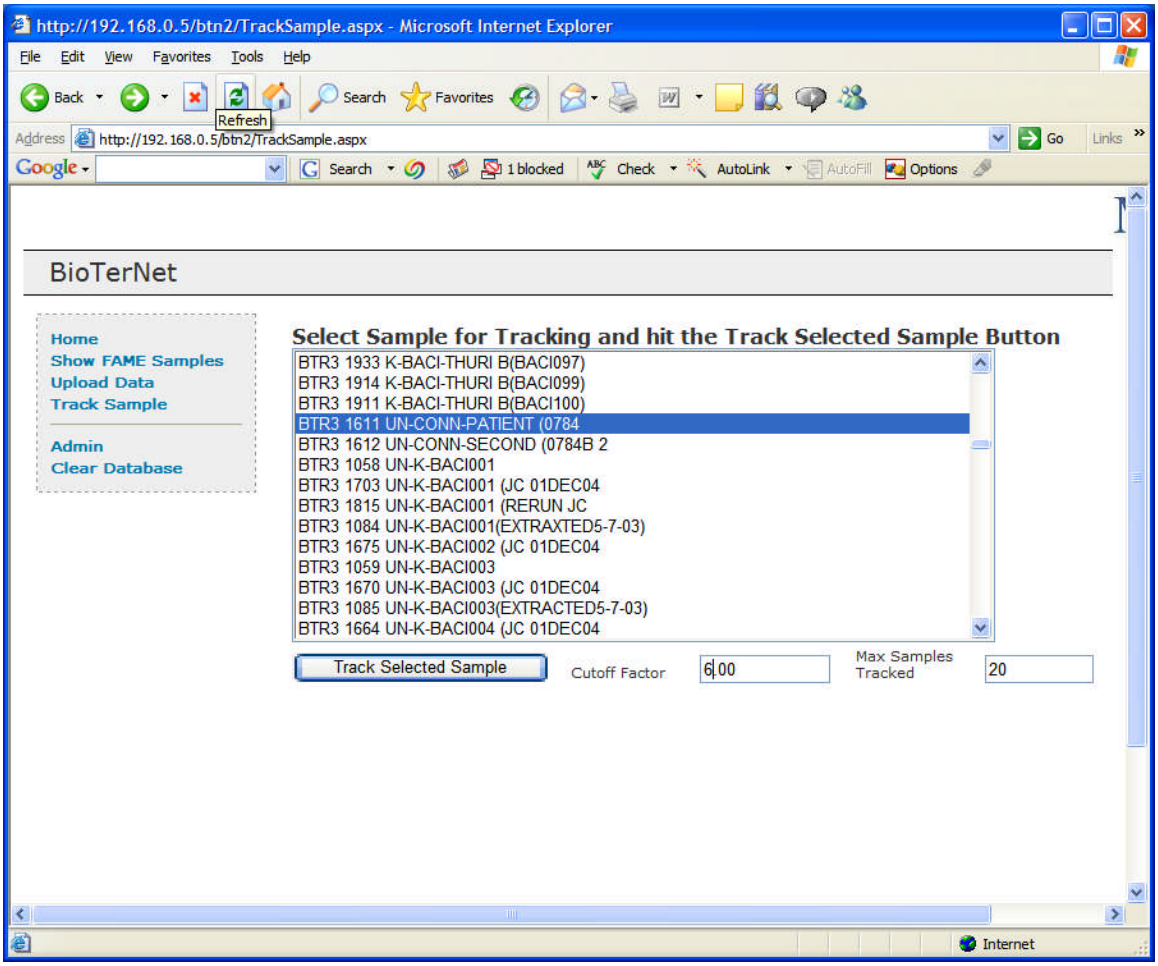
Show Sample Detail

This screen displays bacterial identification information plus the detailed chromatographic information for the selected sample. There is also a link to an additional web site "Names for Life" from Michigan State University that gives detailed taxonomic information for the identified organism.



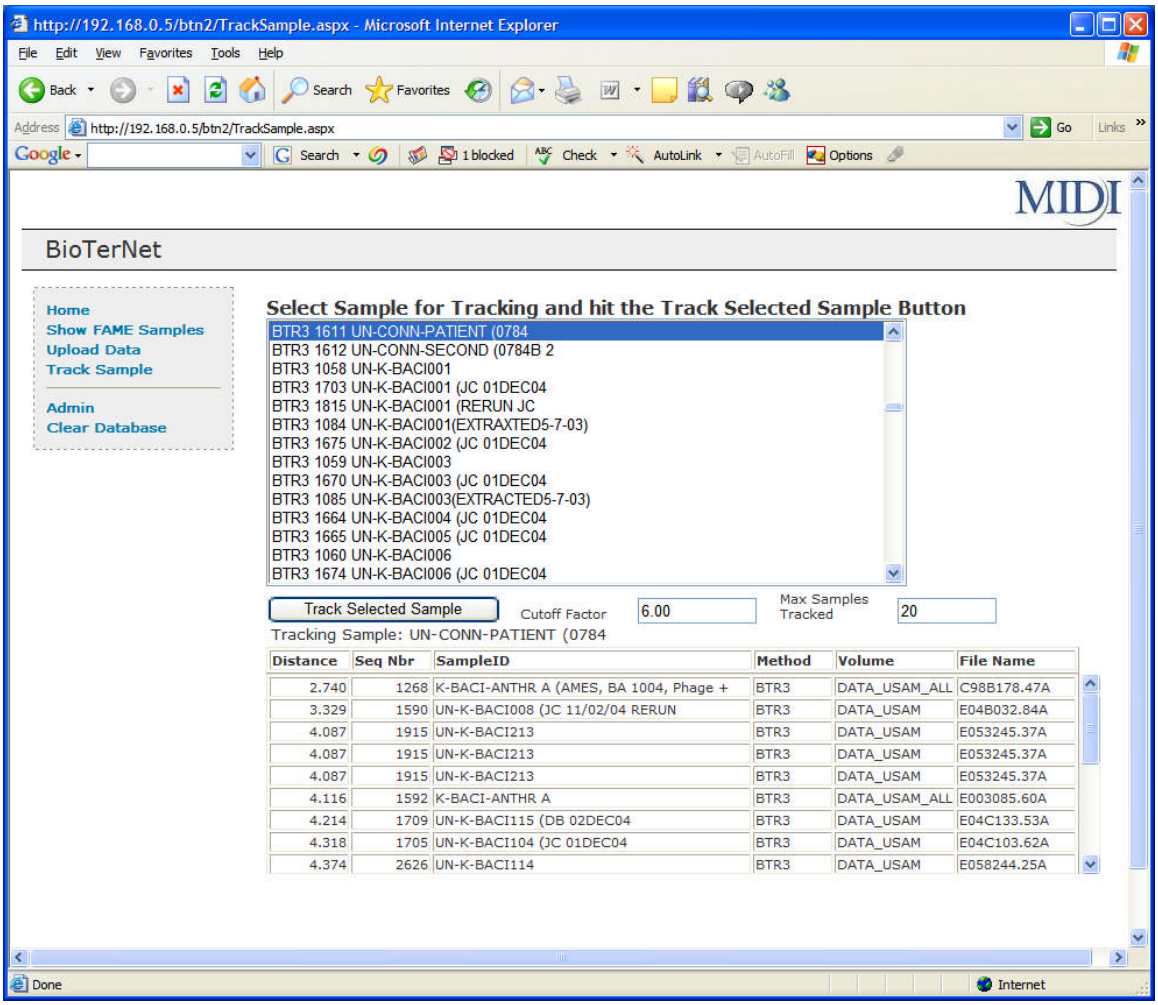
Selecting Sample for Tracking

In this screen the selected sample is highlighted in blue. The cutoff factor is set. The cutoff factor specifies the maximum Euclidian Distance between the selected sample and samples displayed from the database. For a detailed description of the algorithms used to determine how closely samples are related to each other, refer to the Sherlock Tracker Users Manual. The maximum number of samples displayed is also set (Max Samples Tracked).



Sample Tracking Results

The data used in this example consists of anthrax data received from USAMRIID and anthrax data received from the Connecticut Department of Health. This example shows the BioTerNet prototype determining that the case of anthrax that a 94 year old Connecticut woman contracted in 2001 was the Ames strain of anthrax. The best match shown has a Euclidian distance of 2.740 to the run of the Ames Strain of *Bacillus-anthracis* that was run at USAMRIID. The time to make this determination once the data was loaded into the prototype was well under one minute.



Personnel

The following personnel worked on the project during the project year:

Name	*Percent Time Charged*
Mike Alexander	less than 1%
Clarke Arnold	77%
Charles CarterSite	12%
Kevin Clough	4%
Gary Jackoway	43%
Craig Kunitsky	less than 1%
Lindsey Olmstead	79%
Gerard Osterhout	34%
Celia Renai	6%
Myron Sasser	84%
William Stimson	less than 1%

Key Research Accomplishments

- A FAST extraction procedure was developed. This procedure requires only 1 minute per sample compared to the typical 90 to 120 minutes for the current sample preparation procedure. It uses fewer reagents and does not require bulky water baths.
- Optimizing the FAST extraction and the gas chromatographic method for sensitivity, a gain of 750X over the original Sherlock method and 300X over the current Rapid Sherlock method was achieved. This is sufficient for processing single colonies of bacteria from the primary isolation plate and potentially reducing the total time for analysis by 24 hours. More testing and refinement are needed on this.
- Experimental libraries have been created to work with the FAST extraction. Simulation results show that it should have an overall accuracy comparable to the current method's library (85.4% vs. 88.5%). This will improve after accounting for the more reproducible data provided by FAST.
- 784 additional strains of bacteria, covering 539 species, were analyzed for inclusion in the current Sherlock clinical bacteria library. To insure accuracy of their taxonomy, they were analyzed by 16s rRNA sequencing. The enhanced library will be released early in 2007 once the data is fully validated.
- Prototype BioTerNet software was developed and deployed on servers at MIDI. It includes prototypes of database, client/server, and web based application designs that could be used to develop and implement a BioTerNet system. (for instructions on accessing the prototype from the Internet, contact MIDI's software development staff at 302-737-4297).
- Some software and bacteria library advancements were included in Sherlock Version 6.0 and made available to Sherlock users:
    - A clustering algorithm that automates the identification of relationships between samples and helps recognize multiple occurrences of the same strain
    - An updated bioterrorism library with a new *B. anthracis* subgroup and more near neighbors
    - Dozens of entries were added to the total bacteria library set and the taxonomies were updated
    - The ability to include DNA data with fatty acid data and produce a cross-library polyphasic report allows identifications to be made with extremely high confidence
- In January 2006, MIDI received FDA 510(k) clearance for the use of Sherlock for the identification of anthrax (k052485) (see appendix A).

Reportable Outcomes

- Product: MIDI will be able to continue to develop the FAST extraction procedure and associated identification library for Sherlock and release it as an enhancement in a future version of Sherlock.
- Database: Preliminary library of bacterial fatty acid profiles for use with FAST extraction procedure.
- Publication: Clarke Arnold, Craig Kunisky, Myron Sasser, Gary Jackoway and Gerard Osterhout. Department of Homeland Security Evaluation of the MIDI Sherlock Microbial Identification System for the Confirmatory Identification of Bacillus anthracis. International Conference of Emerging Infectious (ICEID 2006), March 2006.

Conclusions

A robust FAST extraction procedure was developed. Preliminary results demonstrate that it will have less variance due to laboratory conditions and operator handling. Reducing the variance introduced by sample preparation, improves the reliability of strain level matching. A sensitivity gain of 300X over the most sensitive currently

available Sherlock method will allow processing single colonies from the primary isolation plate, making the turnaround time 24 hours faster.

More testing is required to confirm these results over a broad selection of bacteria and laboratory conditions. Potential problems with the extremely sensitive method are contaminants and uniform harvesting of a barely visible quantity of cells. Techniques to prevent contamination and tools to help harvest small quantities may be needed. The less sensitive version of the FAST method that was developed and tested does not have these problems. MIDI intends to continue developing the FAST method into a commercial product.

A library for the FAST extraction should be developed. MIDI plans to continue this effort and have a version of the clinical library for organisms with high frequency of occurrence available for beta test in 2006. This will be a valuable tool for clinical labs.

A prototype BioTerNet demonstrated the feasibility of collecting and analyzing epidemiological data from remote sites. The database schema design and implementation used in the prototype will be included in a future version of Sherlock, enabling data to be better organized and analyzed for relationships. The networking and web base client capabilities have value, if there are organizations that are interested in deploying the technology in a network of laboratories. Without a committed client organization or additional funding, the networking capabilities will have a lower priority.

References

Arbeit, R. 1999. Laboratory Procedures for the Epidemiologic Analysis of Microorganisms. In Manual of Clinical Microbiology, 7th Edition, edited by Murray, P.R., E.J. Baron, M.A. Pfaller, F.C. Tenover and R.H. Yolken, 1116-1137. ASM Press: Washington, D.C.

Arnold, C., G. Jackoway, G. Osterhout, and M. Sasser. 2002. Rapid Detection of Nosocomial Infections Using Automated Typing and Data Mining. Abst., Intl. Conf. Emerg. Infect. Dis. 2002, p. 66.

Arnold, C., G. Jackoway, D. Mayo and J. Teska. 2003. Automatic Strain Identification of Bacillus anthracis by fatty acid analysis. Abst. 163, ASM Biodefense Research Meeting.

Arnold, C., and G. Jackoway. 2004. Recognition of Novel Bacteria by Fatty Acid Cluster Analysis. Abst. Q-180, 104th Ann. Mtg. Amer. Soc. Microbiol.

Birnbaum, D., L. Herwaldt, D.E. Low, M. Noble, M. Pfaller, R. Sherertz and A.W. Chow. 1994. Efficacy of Microbial Identification System for Epidemiologic typing of Coagulase-Negative Staphylococci. J. Clin. Microbiol. 32:2113-2119.

Jarvis, W.R. 1996. Selected Aspects of the Socioeconomic Impact of Nosocomial Infections: Morbidity, Mortality, Cost, and Prevention. Infect. Control Hosp. Epidemiol. 17:552-557.

Kim, S.D., L.C. McDonald, W.R. Jarvis, S.K. McAllister, R. Jerris, L.A. Carson and J.M. Miller. 2000. Determining the significance of coagulase-negative staphylococci isolated from blood cultures at a community hospital: a role for species and strain identification. Infect. Control Hosp. Epidemiol. 21:213-217.

Kim, T.Y., J. Woo, T.K. Boon, L.G. Kee and N.L. Ching. 2002. Gas chromatography analysis of fatty acid methyl ester for identification and subtyping of biological warfare agents. 3rd Singapore International Symposium on Protection Against Toxic Substances. 2-6 Dec. 2002.

Leonard, R.B., J. Mayer, M. Sasser, M. Woods, B.R. Mooney, B.G. Brinton, P.L. Newcomb-Gayman and K.C. Carroll. 1995. Comparison of MIDI Sherlock System and Pulsed-Field Gel Electrophoresis in characterizing strains of methicillin-resistant Staphylococcus aureus from a recent hospital outbreak. J. Clin. Microbiol. 33:2723-2727.

Maki, D.G. 1994. Infections Caused by Intravascular Devices Used for Infusion Therapy: Pathogenesis, Prevention, and Management. In Infections Associated with Indwelling Medical Devices, 155-212. ASM Press: Washington, D.C.

Mukwaya, G.M. and D.F. Welch. 1989. Subgrouping of Pseudomonas cepacia by Cellular Fatty Acid Composition. J. Clin. Microbiol. 27:2640-2646.

Nelson, K.E. p. 108-109 in: Infectious Disease Epidemiology: Theory and Practice. 2001. Nelson, K.E., C.M. Williams and N.M.H. Graham, Eds. Aspen Publishers, Gaithersburg, MD. 748p.

Teska, J.D., S.R. Coyne, L.J. Hartman, C.M. Allan, D.R. Christensen, and J.W. Ezzell. 2000. Bacillus anthracis identification based on two automated system databases. Abst. 100th Ann. Mtg. Amer. Soc. Microbiol. p. 186.

Teska, J.D., C.M. Allan, S.L. Redus, S.R. Coyne, and J.W. Ezzell. 2001. Identification of Bacillus anthracis using MIDI whole cell fatty acid analysis. Abst. 4th Intl. Conf. Anthrax. pp. 15-16.

Weinstein, R.A. 1998. Nosocomial Infection Update. Emerg. Infect. Dis. 4(3): 416-420.

Appendix A

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
2098 Gaither Road
Rockville MD 20850

JAN 17 2006

Mr. William Stimson
Quality Assurance/Compliance Manager
MIDI Inc.
125 Sandy Drive
Newark, DE 19713

Re: k052485
Trade/Device Name: MIDI Sherlock Microbial Identification System BioDef™
Regulation Number: Unclassified
Product Code: NWZ, KZQ, NPO
Dated: December 8, 2005
Received: December 9, 2005

Dear Mr. Stimson:

We have reviewed your Section 510(k) premarket notification of intent to market the device referenced above and have determined the device is substantially equivalent (for the indications for use stated in the enclosure) to legally marketed predicate devices marketed in interstate commerce prior to May 28, 1976, the enactment date of the Medical Device Amendments, or to devices that have been reclassified in accordance with the provisions of the Federal Food, Drug, and Cosmetic Act (Act) that do not require approval of a premarket approval application (PMA). You may, therefore, market the device, subject to the general controls provisions of the Act. The general controls provisions of the Act include requirements for annual registration, listing of devices, good manufacturing practice, labeling, and prohibitions against misbranding and adulteration.

If your device is classified (see above) into either class II (Special Controls) or class III (PMA), it may be subject to such additional controls. Existing major regulations affecting your device can be found in Title 21, Code of Federal Regulations (CFR), Parts 800 to 895. In addition, FDA may publish further announcements concerning your device in the Federal Register.

Please be advised that FDA's issuance of a substantial equivalence determination does not mean that FDA has made a determination that your device complies with other requirements of the Act or any Federal statutes and regulations administered by other Federal agencies. You must comply with all the Act's requirements, including, but not limited to: registration and listing (21 CFR Part 807); labeling (21 CFR Parts 801 and 809); and good manufacturing practice requirements as set forth in the quality systems (QS) regulation (21 CFR Part 820).

Page 2 –

This letter will allow you to begin marketing your device as described in your Section 510(k) premarket notification. The FDA finding of substantial equivalence of your device to a legally marketed predicate device results in a classification for your device and thus, permits your device to proceed to the market.

If you desire specific information about the application of labeling requirements to your device, or questions on the promotion and advertising of your device, please contact the Office of *In Vitro* Diagnostic Device Evaluation and Safety at (240)276-0484. Also, please note the regulation entitled, "Misbranding by reference to premarket notification" (21CFR Part 807.97). You may obtain other general information on your responsibilities under the Act from the Division of Small Manufacturers, International and Consumer Assistance at its toll-free number (800) 638-2041 or (301) 443-6597 or at its Internet address http://www.fda.gov/cdrh/industry/support/index.html

Sincerely yours,

Sally A. Hojvat, M.Sc., Ph.D.
Director
Division of Microbiology Devices
Office of *In Vitro* Diagnostic Device
Evaluation and Safety
Center for Devices and
Radiological Health

Enclosure

Indications for Use

510(k) Number (if known): k052485

Device Name: MIDI Sherlock Microbial Identification System BioDef™

Indications For Use:

The MIDI Sherlock Microbial Identification System for BioDef™ is intended to aid in the identification of *Bacillus anthracis* through the analysis of membrane fatty acids derived from cultured bacterial samples, using gas chromatography and pattern recognition software.

Warnings:

- The Sherlock BioDef system is not intended for use with spore preparations or materials other than colonies from trypticase soy agar (with 5% defibrinated sheep blood) that have been presumptively identified as *Bacillus* spp. BioDef cannot assess presence or absence of virulence factors.
- Accurate identification of *B. anthracis* is dependent on purity of isolates, length of incubation (24 h) at 35° C, and media type, along with well-trained and experienced technicians. Only BBL™ 5% sheep blood agar has been validated for identifying *B. anthracis* with the MIDI Sherlock Microbial Identification System for BioDef™.
- The definitive identification of *B. anthracis* from colony growth requires additional testing and confirmation procedures in consultation with public health or other authorities to whom reporting is required.
- Identification of organisms other than *B. anthracis* in the Sherlock BioDef's system's BTR library database has not been evaluated.

Prescription Use ✓ (Part 21 CFR 801 Subpart D) AND/OR Over-The-Counter Use _____ (21 CFR 807 Subpart C)

(PLEASE DO NOT WRITE BELOW THIS LINE-CONTINUE ON ANOTHER PAGE IF NEEDED)

Concurrence of CDRH, Office of In Vitro Diagnostic Devices (OIVD)

Division Sign-Off

Page 1 of 1

Office of In Vitro Diagnostic Device Evaluation and Safety

510(k) K052485